The Dissertation Committee for Michael Duane Bradshaw
Certifies that this is the approved version of the following dissertation:

A Case Study: Views of Bermudian stakeholders in education on the Accountability of an Outcomes-based Accountability System.

Committee:

William Moore, Jr., Supervisor

John E. Roueche

Norvell Northcutt

James Scheurich

Uttarayan Bagchi

A Case Study: Views of Bermudian stakeholders in education on the Accountability of an Outcomes-based Accountability System.

by

Michael Duane Bradshaw B.Sc.; M. Phil.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Education

The University of Texas at Austin

December 2003

Dedication

The question of accountability is especially meaningful to me because of a particular world view and personal political philosophy that have been burnished by being a male and a person of colour. Today, these offer contradictory expectations of power and therefore of one's position in the hierarchy of accountability. My grandmother provided my first education --- both the formal learning of the alphabet and the like, but more importantly, a sense of self and of mores and values and the responsibility of power. I therefore applaud Dr. John Roueche for a long pursuit of formal institutional and organizational accountability in education. Dr. Moore is acclaimed for his focus on acceptance and exercise of accountability at the individual level. I recognize Drs. Bagchi, Northcutt and Scheurich for their efforts in improving my tool repertoire and my understanding of accountability. Mostly, I dedicate this to my sons, Kumi and Kieran, by whom I expect to be held accountable, and to all grandsons yet unborn, who shall inherit the legacy of our efforts.

Accountability that yields universal access and success in education will deliver a world that is fair and effective and efficient.

Acknowledgements

I must acknowledge the Lord from whom what little I have been able to offer in this effort is but a small gift of his mercy and bounty. I thank Him as I have been kept in the necessary physical and mental health to complete this project.

My wife, Janet, and my family have supported me throughout this undertaking. Many of the people of Bermuda have assisted and made themselves available in whatsoever ways that were appropriate. Numerous persons at the University of Texas at Austin have been critical in enabling me to continue my education whilst still at a great distance from Austin. It may sometimes have appeared to be a part of an assigned job or responsibility of a post but I can attest to efforts and attitudes that reflect a sense of self accountability to the customer and a desire to exceed the customer's expectations.

Additionally, I note those persons who contributed very directly to my efforts by their direct participation. This includes colleagues who assisted with the pilot in Austin. Many thanks. There follows the Bermudian respondents who, by their efforts, have launched me on a life's quest. I acknowledge of course my Dissertation Committee Chair and Members. I trust that I will be able to show my appreciation to them by the quality of my further efforts in the field.

A Case Study: Views of Bermudian stakeholders in education on the Accountability of an Outcomes-based Accountability System

Publication No._____

Michael Duane Bradshaw, Ed. D.
The University of Texas at Austin, 2003

Supervisor: William Moore, Jr.

In answering the accountability mandate and competition from the private school/charter school system, public education authorities have introduced various means to measure the performance of the public school system so as to improve it. In this case study of the Bermuda public school system, using self-administered questionnaires, respondents provided personal demographic information and their ratings of comments or determinants/criteria used in outcomes-based accountability systems. Analysis of the data showed that educators and the general public had much general similarity in ratings of the major attributes that were important for accountability considerations but were still distinctive. Using factor analysis of the data and then ANOVA with demographic scores as Independent Variables and the Factor Scores as the Dependent Variables suggests some significant relationships can be found and these differ for the Educators and the Public.

Considering the stakeholders' view, educational accountability requires Total Quality Management (TQM) to be a management style of commitment of the

total resources of the organization to meeting the customers' satisfaction and even exceeding their expectations. Despite much public experience with it, TQM's impact on education has been limited. The pervasiveness of the accountability mandate in higher education and concerns about the comparative quality of public and private K-12 education underscore the need to go beyond a survey of the public's attitudes about the public school performance as determined by the PDK/Gallup poll.

Using this case study the researcher produced and tested a survey tool to gather data for testing differences in opinions and clarifying which criteria/indicators were influential in these opinions. The study increases the education literature on Bermuda. Since Bermuda's accountability environment has some parallels with that of the USA as exemplified by Texas, the research project adds information about how we might review and reflect on such outcomes-based accountability systems as are set up in Bermuda and in Texas. This is significant for the public's involvement in centralized and de-centralized education. This methodology may be valuable in such related research.

Table of Contents

Chapter 5 Discussion of Analysis and Interpretation of Results

Chapter 6 Significance of the study

Appendices

List of Tables

Chapter 1 – Introduction and Development of the Research Area

Introduction to accountability

Accountability refers to the act or process or report that explains or takes responsibility for standards being met or not being met. Education has always been associated with accountability. Students demonstrate their accountability for what they have learned by taking tests and quizzes where success in meeting the standard is rewarded (by stars or letter grades or GPA scores, etc) and failure to reach the levels for which they are held accountable will warrant sanctions (no stars or undesirable grades and low GPA scores). Early American college students were subjected to rigorous examination via the senior declamation at the end of baccalaureate studies. National testing systems appeared in higher education in the 1940's in the form of the Graduate Records Office of the Carnegie Foundation (producing the Tests of General Education), The Co-operative Test Service (administering the General Culture Tests) and then the Educational Testing Service. Alverno College, from the 1970's, has been an innovator in research in college-based assessment programmes and has an excellent reputation for emphasizing the intellectual and personal development of its students. Individual efforts have thrived elsewhere but there has only been limited effort at a national response for tertiary education accountability.

Measuring student achievement has become a central policy issue of higher education. Colleges and universities are themselves assessed by the six regional accrediting agencies that measure overall institutional effectiveness. Nationwide, almost 200 different accrediting bodies examine specialized courses and programmes on the campuses in areas from agriculture through nursing and on to welding. If courses or programmes do not pass scrutiny, then

graduates either may not sit the licensing examinations or will not receive other criteria that are required for graduates to practice in the area studied. At the university level the tenure system and other self-regulatory features hold the faculty accountable. Accountability exists everywhere in our everyday world. Yet, we have been reluctant to apply the full rigour of accountability throughout education.

Accountability systems are designed to maximize excellence in the product or service and to maximize the added value for the invested resources. This description shows the ready identification that exists with business models. The operation and practice of accountability undergird today's better living standards. In the hinterland or metropolis, in a remote Third World village or a Central Park condo, the lifestyle is one of ease and luxury compared to the normal expectations of a few generations ago. For this some thank the Druckers and Demings of the business world and their strategies of Management By Objectives (MBO), of Total Quality Management (TQM), of Continuous Quality Improvement (CQI), of Six Sigma and other accountability. ("The enigma of six sigma" 2000; "Motorola six sigma" 2000; Community College Journal of Research and Practices 1994). Despite stringent accountability regimes being applied in manufacturing and service industries, to the military and to medicine, education steadfastly resists.

Education depends upon those who work directly with the student in the business of teaching-learning; i.e., the educators, and those others who manage the system in which education occurs, i.e. the administrators. Many educators and administrators insist that education is not an industry and students are not products. Nor is education like any other

service industry. Hospitals and hotels also provide a service and these have bowed to the accountability drive but some claim that education is too complex to be measured accurately and so it is better not to measure it at all. Generally, unions in education resist the use of merit pay because the conventional wisdom is that every teacher is doing a good job and so all should be rewarded (Parker & Slaughter 1994; The 1998-'99 National Association of College and University Business Officer Benchmark Program). Besides, one class may be predisposed to success because of the parents and the home environment and the culture (whether the teacher gives extra effort or not) whereas another class lacks these starting advantages and so does not achieve at the same level (despite the teacher's outstanding extra efforts). They claim that it would be unfair to reward the first teacher and not the second. Thus there are problems with the lack of precision and fairness when accountability is applied broadly to the education scene. Yet that is the rub. Students are held accountable despite these problems. It must also be acknowledged that educators and administrators may fail to do an adequate job just as may surgeons and accountants. The public is insisting upon accountability in the education systems. Citizens want all parties to be accountable and not just the students.

Accountability and the History of US Education

Accountability has been particularly difficult in the USA because of the very peculiar mechanism of decentralized education (Vinovskis 1998; Black et al 1987). Education has been carried out in individual districts in individual states with nothing even close to a national system until very recent times (The National Education Goals Report 1999; Mitchell, 1996). The responsibility for education as the right of each state with no responsibility to a central

authority is a basic tenet of federalism and some states put more into education than others (Black and Black, 1987). Since it is the tax base of the school district that funds education and determines the quality of schooling available, then poor regions get poor schooling. The poor had very little political voice and were even more oppressed in former times. They made no real demands of accountability. In general, even enlightened liberals thought that to bring the working class up to middle class standards and values was a noble gesture and so the methods of education that worked for the middle class were surely appropriate for all. In this time when it was a new idea to educate women and Black people and the education of the poor was barely accepted, it was unthinkable to make too great a demand about the efficacy of education. Traditionally, accountability in public schools had been limited to fiscal correctness and inputs. There were checks as to what was bought and how wisely the money was spent but there was no real evaluation as to whether or not the job of teaching-learning was being effectively pursued. As long as the correct number of books was in the library; or the correct proportions of credentialed personnel existed on staff, or the designated time was allocated for subject areas --- the inputs were correct, then it was assumed that was all that was needed for success.

There was no real debate about what the role of education was? Until the 1950's education was segregated and not universally available. Attendance was not always compulsory and there was often no tracking of requirements for class promotion or graduation. Qualifications and competence of educators and administrators could vary wildly. The Education Commission of the States started in the mid-60's as a national policy making and planning collaborative but each state still made independent policy decisions. The Elementary and

Secondary Education Act of 1965 opened up education to all. The real emphasis on quality came later. Frightened by the Russian's launch of Sputnik in 1957, Americans selectively expanded and supported higher education and some junior science areas. It was difficult to compare American education to that of other countries since reliable and consistent statistics did not exist at the national level for the USA. Still, decades later the studies that led to publication of "A nation at risk." in 1983 highlighted a dangerous, poorly functioning system. The School Improvement Act of 1987 and then the National Education Goals Panel of 1990 (NEGP) were serious efforts to remedy the situation. The NEGP emerged from the Charlottesville Summit, which was the first time in USA history that federal and state political leaders and professionals from both parties met with the President of the USA and reached consensus on specified national education goals and priorities, and set timelines and outlined strategies to achieve them by 2000. The NEGP included a requirement for a forceful accountability mandate on public education.

Accountability in the present public education system

It is business and the upper middle class who drive accountability. The accountability systems used in business seek to have fewer people each year produce more items at a lower cost than the year before. Such escalating productivity has depended on technological progress but the limiting factor is fast becoming the ability of the workforce to understand and implement that technology. Even when a high school education was the passport to a solid start for a middle-class life, a number of persons were deficient in the basic skills. But since the norm was not that high and few persons comparatively went on to higher education, these shortcomings did not really stand out. Persons improved on the job; poor readers (after graduation) could

become better with time. America was largely agrarian until the middle of the twentieth century and it is interesting to see how states like Florida and Texas changed after WWII from a farm base to a manufacturing and industrial economy. Entrepreneurs who wanted skilled labour at low rates therefore pushed some of the early reforms in education. This meant a broadened access to education (at least at the lower levels). Improved education was to be an economic engine that would pull the state economy along in the new circumstances. Education of the masses was to allow forward-looking entrepreneurs to have enough of a prepared and low cost workforce to attract industry. This is still being seen in various Southern states today (Personal communication from Dr. Cameron; Dr. Wallin; Dr. Hodges; Dr. Wallace).

The first real success of the Civil Rights Movement is usually celebrated as the decision of the Courts in Brown v Board of Education in 1954 to broaden access to education. The final outlawing of segregation in public schools occurred by the Civil Rights Act of 1964. The Civil Rights banner has been a powerful lever that has advanced social justice for all other sectors as well. Whereas the infamous Dredd Scott decision of 1857 had not recognized the existence or equity of Negroes under the American Constitution, the Civil Rights Legislation repudiated this inequity. Lawrence Fuchs (1990) in "The American Kaleidoscope: Race, ethnicity and the civic culture." noted how Black men and women discarded "the psychological stigma of powerlessness inflicted on them by caste for generations" and became empowered and emboldened to claim their rightful place and an equal place at the table of civic culture. Using this step taken by African-Americans, all the other deprived groups of the American cultural fabric (whether based on race and ethnicity or gender or sexual orientation or disability or religion) claimed the right to hold society accountable for their own access and

equity issues in education. Title IX and the Women's Education Act of 1974 are examples of affirmative action that addressed the institutionalised, systemic and long-existing disadvantages under which those other than non-White males had sought education. The public now expected and sometimes demanded access to quality education for all. Education was to be an economic driver and as economic enfranchisement occurred greater political enfranchisement would follow and the weave of the civic fabric would become even tighter. It is therefore important that the public holds education accountable and is able to be satisfied in its needs, desires and expectations.

Accountability in education in the first decade of the twenty-first century

The clash of institutional autonomy versus political accountability has always been at the heart of arguments to avoid the regulation of education. Education was to be controlled by the local populace as they should be the best judge of what was suitable for their needs. Furthermore, the university lecturers had won the right to intellectual freedom that allowed them to be unassailable by administrators in most matters that concerned the classroom. Lecturers vigorously defended and operated their own self-regulation and accountability to the standards of their discipline and to peers through the tenure system. Classroom teachers were under far more authoritarian control and have been held accountable in many direct ways to the principal, the school superintendent, the school board and even influential citizens. The community college faculty forms an interesting middle ground. The first administrators of the community colleges often came from the ranks of superintendents and secondary school principals and state education officers. Both faculty and administrators were usually the products of middle tier colleges and universities. Thus the community college, which was

begun as a junior college movement to carry out the work of the early undergraduate years of the senior 4-year institutions, held persons with mixed experiences. Many wanted the intellectual status and monitoring of self as seen in the senior units but they did not accept the concomitant responsibilities of self-regulation. Others are comfortable with the system seen in the K-12 level where there is an adversarial type approach between administrators and educators in respect of accountability issues.

The K-12 system is the foundation for all higher levels. Accountability at any level is therefore predicated upon success from accountability at K-12. The overarching problem is that none of the current stakeholders is more than moderately satisfied with the quality delivered by the present public education system at the K-12 level. The situation in Texas provides a useful guidepost and barometer for what is happening nationally as it has seen recent gains that have been attributed to the state accountability system. As recent as 2000 Finn and Petrilli reported "that 42 states still hold mediocre or inferior expectations for their K-12 students, at least in most subjects. Hence it must be said, 17 years after "*A* Nation at Risk", 11years after the Charlottesville Summit, and in the same year that our 'National Education Goals' were to be met: most states still have not completed the first step of standards-based reform." This is confirmation of the great disparities in educational opportunities and success that are displayed in Jonathan Kozol's book 'Savage Inequalities: Children in America's Schools' (1991). Employers and higher education point at the public K-12 system and complain loudly that a good job is not being done. Consequently, statehouse politicians have mandated detailed accountability systems.

Although all 50 states claimed to have accountability systems for public education, Texas was one of only 5 that had both a strong accountability system and solid standards against which performance was measured. Texas ranked 3rd in the nation in 2000 and showed a slight fall from its former position of 2nd in 1998. Such progress on the education front has been by a purposeful and intentional strategy as the Texas legislature has provided sustained financial and logistical support and a forceful accountability system. There is an expectation of uniform standards of performance in all subjects for all students and providing disaggregated results according to race/ethnicity and gender monitors this. The Mission Statement of the Texas Education Authority (TEA) "is to ensure that all Texas children have access to quality education that enables them to achieve their full potential and fully participate now and in the future in the social, economic, and educational opportunities of our state and nation." (State Board of Texas, 1995).

The use of a state-wide criterion-referenced assessment system (Texas Assessment of Academic Skills or TAAS) and a state-wide curriculum to measure performance at Grades 4, 8 and 10 are integrated in the Texas accountability system. TAAS began in 1990 and has been stable in content and format since the mid-1990's. At least 5 years of consistent implementation has been available and this is more than for many other jurisdictions. There has been sustained increase both in the numbers of persons taking TAAS and the overall performance. Haney (2000) has criticized the avowed link of the high accountability system to the extraordinarily large gains in Reading and Mathematics scores on TAAS for non-White groups. These results have not been substantiated by independent testing (Bello, 2000;

Wilgoren, 2000; Yardley, 2000) and chicanery has been revealed in at least one region (Haney 2000; Klein et al 2001; Miller 2001).

The TAAS was not the sole criterion for holding schools accountable. TEA officers used a state-wide, comprehensive student-level information system to collect and collate 87 different data inputs that contributed information in 9 major categories other than the TAAS results. The standards for each rating category are set within guidelines specified by the legislature. Standards were designed to phase in ever increasing expectations of performance at the district and even campus level. Thus requirements for 'acceptable' ratings have been raised each year since 1996. Relevant personnel in local districts were encouraged to develop and operate local accountability systems that complemented the state-wide system. District and school ratings that were based primarily on TAAS scores, Dropout Rate and Attendance Rate to decide all rating levels were published. Other indicators can be used for assessing performance excellence, rewards and reports. One such report is the School Report Card. This is used to meet a statutory obligation that every school must provide each student's family with a copy of its complete School Report Card showing criteria used for the school's rating as well as further information. Indicators were used for a number of years first with gradually extending of those used as the base indicators for rating.

Each campus and district was assigned a rating from a range of 4 levels from Exemplary to Academically Unacceptable/Low Performing. To recognize the diversity of situations Special Circumstances Ratings were employed. Schools or districts that had problems with meeting minimum targets or improvement were assisted by TEA staff and received increased

intervention if the situation persisted. The Legislature would also fund the transfer of a student from a chronic, low-performing school to another of higher performance results, even if outside the usual district boundaries, at the parent's request. On the other hand monetary rewards were available to school campuses where there was performance or improvement that was judged worthy.

Thus we have the situation that the public and the business community expressed dissatisfaction with the performance of the then education system. In the absence of any appropriately substantive response by the public education system the Legislature instructed the Texas Education Agency and the Texas State Board of Education (SBOE), as the responsible agencies for guiding and monitoring all activities of K-12 public education in Texas, to set up the accountability system. The SBOE has allowed the formation of some 170 charter schools as additional alternatives to the public schools and all are operated under the same accountability criteria. From an education service centre in each of the 20 regions, staff members provide services that enhance student and school improvement. Other associations/organizations such as the Charles Dana Center (U. T. Austin) may collaborate to enable further educational excellence to Texas.

The school campus and the district are held strictly accountable for meeting performances as compared against clear and known standards. The accountability system is imbedded in law and operated by formal procedures that are both stable and consistent but allowing of growth and progression in a predictable and agreed fashion. Sanctions and rewards are appropriately instituted for non-compliance or exceeding compliance with the standards.

The Research Area

Texas has a clearly defined accountability system for K-12 but the perspective on the performance of the public education system is more muddled. There is doubtless some improvement over the segregated system of earlier days but there are still sharp discrepancies in public education in various parts of the state. There are still schools that have very poor ratings on the accountability system. The belief is however that in time the application of the accountability system will lead to improvement for all. It is therefore instructive to consider the Texas accountability system for K-12 public education and the performance of the K-12 system. As in many other states public dissatisfaction with public K-12 education has led to a growing private system of charter schools. The private system is smaller and has produced better results for its range of students and by extension it is thought that the private system may be a better system in all respects. Texas does not have a personal state income tax and the funding for education is mainly generated at the local level from property taxes. Some equalization is attempted by the legislature requiring the very wealthy districts (about 10%) to assist their less well-funded colleagues in various ways. This is nicknamed 'The Robin Hood Law' in education circles. Texans also have a strongly individualistic and independent political culture, which favours maximal local control. Thus some stakeholders are more visible and voluble than others inn the exercise of accountability. Accountability to a central authority is tolerated rather than readily appreciated or valued.

The research area concerns attempts to modify the structure of the accountability system so as to enhance the general improvement of the performance of the K-12 public education system. Schemes that manage accountability are intended to lead to improvements in the quality of the

service delivered and increases in productivity. In a comprehensive study of school districts that have shown that substantial and sustained improvement can be achieved by 'economically disadvantaged students', Scheurich et al (2001) reported that school board members, school administrators and teachers all attributed the substantial impact of an outcomes based accountability system with public responsibility as a driving force for changes in attitude and behaviour in the four districts studied. Fundamental to the success was that personnel in these districts held themselves and each other accountable for student success. "All four districts had thus developed practices for supervising, evaluating, and holding people accountable for practices that contributed positively to the instructional mission of the district." They further noted that "Each study district, however, had developed specific, local accountability practices that ensured that everyone involved in the instructional program was held accountable for educational equity and student achievement."

The Specific Problem Area

This project examined a noticeable deficit area that is seen in the operation of several education accountability systems. The Texas system was designed with very little input from the general public and even from educators. It was developed by state education officials under a political mandate that "called for a system of accountability based on student performance" (Texas Education Agency, 2000). There is much academic debate about that system's success but no one appeared to be asking the various stakeholders what is their view on the accountability of the accountability system.

Answers to a number of questions were sought in order to explore and illuminate the research area. A first question was --- What is the opinion and satisfaction level of stakeholders as regards the success of the outcomes-based accountability system? What criteria can stakeholders identify that will allow them to evaluate the working of the accountability system in a more concerted manner so as to lead to further improvements? Also, what are the indicators and measures of effectiveness of performance of the K-12 education system that can be designed and implemented from such criteria?

Hypothesis

If stakeholders in public education, especially those other than education professionals, are given the broadest opportunity to contribute to the content and assessment of the education accountability system, then they would actively participate and an improved product would result. Sclafani (2001) reported that, "The public nature of the data from the accountability system, and especially the disaggregated data, made clear where schools were." When the public was aware "the school rallied to solve the problem." Quality management systems posit that the customers hold the key data for their satisfaction and the success of the enterprise (Motorola Six Sigma, 2000). In his study of districts that showed marked performance improvement consonant with the application of the Texas accountability system, Scheurich (2001) noted the association of dramatic changes in the attitudes and behaviours of administrators and educators with "catalytic events or circumstances" that included public dissension by community activists or by business leaders or by school board members of the data revealed by the accountability system. Stakeholders will act if circumstances are extreme and it remained to be seen if they would respond before a crisis was at hand.

Definition of terms

(Accountability) Indicators	specific measures/standards of the accountability criteria.
(Accountability) criteria	specific items on which accountability judgements are based.
Accountability mandate	an electoral imperative to implement accountability.
Academic needs	competencies of computation, communication in writing and speech that are the basics of education.
Accountability ratings	categories denoting the level of performance as per accountability criteria.
Accountability system	the set of rules, procedures and policies by which the item is evaluated as having met or not met acceptable benchmarks.
CQI	management scheme that seeks incessant incremental improvements in all areas.
Charter schools	state-licensed schools that are privately funded/operated.
Criterion –referenced	evaluated against set criteria rather than another's performance.
Educational desires	specific content knowledge and skills that may be desired to fulfill a particular ambition or goal.
Educational expectations	specific minimum skills and competencies such as problem solving and self-directed learning that are held to be a basic.
K-12 public education	state-provided education from kindergarten to form 12.
MBO	management methodology that focused on having clear objectives that always were to be kept to the fore.
Six Sigma	management theory that uses mathematical modelling of any enterprise to promote sharp turn-arounds in quality.

Stakeholders	any person or group of persons who can derive direct or indirect benefit from the operation of the education system.
Standardized testing	a system of academic tests that are designed to be unbiased and which are given to all students at the same time.
TAAS	standardized tests in reading, writing and mathematics that are taken state wide by students at certain grade levels.
TQM	management theory that uses the pursuit of quality in all performances and views the customer's satisfaction as being paramount.

Discussion of Assumptions

A number of assumptions were made. The first was that accountability procedures and mechanisms that were commonly applicable and successful in the world of work and business would also hold in education systems. Secondly, as accountability systems in education are rather recent it means that they may undergo serious future modifications necessary. Thirdly, some sort of formal accountability of education is here to stay.

Finally, stakeholders would exercise sufficient objectivity that the research would not degenerate along the Special Case Exclusionary Principle. This would say that the system as designated was perfectly sound and should be fully supported except for a number of 'special cases which should be excluded' from having to be bound by the constraints of the system that apply to all other cases. These special cases would be identified during field observations and experiences. This would be a recipe for disaster as the product would be so vague as to be ineffective. The operations of school boards and other units show that if the mechanism is

appropriate and facilitators maintain some distance between the individual's view of individual situations, then people are induced to act in the best interests of the group and less noticeably in their own self-centred interests. Hard decisions can be made from an objective perspective.

Limits and cautions

Transferability of principles from the business world to education.

There is some concern about the use of methodologies and strategies that have been successful in the corporate sphere but which are unproven by practitioners in education.

Use of Bermuda as a study system

A problem of using a single study system (Bermuda) is that any uniqueness may limit the extent to which it can give general information about American accountability systems.

Possible hesitant responsiveness and confidence of respondents

The lack of confidence and prior experience by respondents could limit responsiveness. Lay persons seldom can formally comment about ways of holding the education system accountable. Educators have been notoriously restrictive about external evaluation and foster a general belief that only professional educators can properly evaluate education.

Limited previous data in the area

Few other examples of using the general population to secure information on the accountability of education were located. There was not a ready model for customer-centric research in education or against which findings could be evaluated.

Significance of the study

The study was an empirical attempt in using business type methodologies to pursue accountability in the education sphere. This is becoming more relevant as more partnerships are established between the two areas. Not only do business persons serve on governing bodies in K-12 and higher education but working collaborations are being established as education is seen as an important asset to social and economic progress and stability. Lay persons should certainly want a better understanding of accountability issues as they apply to education, and educators and administrators should welcome the contributions of their business partners. This study explored commonalities and distinctions of accountability methodologies in business and education. It will be of value to all stakeholders in education in Bermuda and may be able to be extrapolated elsewhere.

Accountability is necessary for active and democratic local control. The public can hold education accountable if they have the means and the confidence to do this in an effective and appropriate manner. This study provided them the chance to show such confidence and to explore the necessary means. Hence the benefit to the good governance of educational bodies is clear. A second benefit from enhanced accountability would be the increased competitive status of education. With continuous increase in the bidders for public funds it will certainly be an advantage for the education system to show itself as being accountable in a way that has broad acceptability within the community. Accountability also fosters excellence. This project displays enhanced accountability in the public K-12 system and therefore presumably encourages excellence. Such features should be significant to citizens, educators and administrators in the Bermuda and Texas K-12 systems. The Texas K-12 system shows

evidence of particular attention to being upgraded and to emphasising accountability in education.

Almost every individual has experience of K-12 education and has an opinion about its performance. The public K-12 system is centrally important in the lives of individuals and in society as the basis for all later education. It involves and affects substantial quantities of human, financial and other resources. Policy and research units are therefore very concerned about accountability of the education system and how this is viewed/analysed by the public. Texas has a scheme of centralized planning by the Texas Education Agency and the State Board of Education and of local control by school district boards. Bermuda has central education planning by its Ministry of Education with modest local control by school boards. Academics, central planners and others in policy and research in education and affected areas should find this research useful. The study advances a methodology for public involvement that is formal and that has an academically defensible base.

Special attention is drawn to the significance that exists for the community college world. Accountability has been intimated for higher education and a number of schemes are in exploratory stages. The regional bodies have become increasingly insistent about including institutional effectiveness measures and standards as indicators of performance for the community colleges for accreditation purposes. Community college educators and administrators are of two mindsets generally. Those with experience of vocational/career track programmes are very conversant with external bodies that hold the institutions, individual programmes and finally the students accountable. Conversely, those who favour autonomous

self-regulation like the 4-year colleges and universities steadfastly refuse and fear any accountability other than self-accountability. As public accountability demands increase in the future the community college would do well to have learned from the K-12 system. The community college world will definitely benefit from seeing how different stakeholder interests are explored and addressed in the public K-12 system.

The literature relating American education at the lower levels to that of other countries is small and has only begun to grow recently. American K-12 education is decentralized and with the intentional lack of a national curriculum or national reporting system there has been no means of providing consistent and uniform statistics (National Center for Education Statistics, 2000). Notably, one requirement of the Accountability Mandate from the Governors Conference of 1986 was that all states would keep statistics that were comparable (Vinovskis, 1998). In the various states different accountability systems are being developed and operated. These research findings contribute to the literature and may partly copy the involvement of the public in school accountability in Europe where school rankings may be published like league tables showing performance of sporting teams.

Research questions

The research questions sought to have stakeholders identify their academic needs, desires and expectations of the educational system themselves. This contrasted to the usual method that involved professionals interpreting and anticipating what would be the needs, desires and expectations of various stakeholders. The questions were as follows: -

What are the opinions of Educators and the Public in Bermuda about an outcomes-based accountability system for public K-12 education?

What criteria will the various stakeholders identify as being appropriate and important to allow them to better manage and influence such an accountability system for public K-12 education?

How do the criteria identified by these various stakeholders differ and how are they similar?

Chapter 2 – Literature Review

"Systemic reformers envision the rapid creation of a system marked by strong and consistent standards, coherent guidance for instruction, strong consensus about goals and much greater equality in educational achievement." David K Cohen , in 'Holding Schools Accountable: Performance-Based Reform in Education. p.101. (Ladd, H. 1996)

Overview of Chapter

Accountability has been a watchword for a long time in education without much really definitive progress. The current meaning originates from the business world. This is where most persons have encountered accountability and they bring this into the rest of their experiences as taxpayers, as school board members, as parents and students. A survey of recent and current accountability models in the business world is made to indicate their benefits. The transfer of such principles of accountability to the education arena is discussed with a general review of examples. Special focus is centered on various state systems for K-12 education accountability. This illustrates America's opportunity to simultaneously develop and evaluate several solutions to a problem because of the autonomy present in a decentralized situation. The Texas system is detailed as an example of a benchmark accountability system.

Accountability is a persistent consideration in higher education and accountability mechanisms are commonly emphasized in the community college as "Institutional Effectiveness Measures". Community colleges offer a very good chance for successful application of accountability strategies from the business world. They generally have more interaction with the business world (through many partnerships types) than either K-12 units or the senior universities. A sample of accountability systems in the non-American context is also explored to illustrate contrasts and similarities between the very decentralized American system and the typically centralized non-American systems. Researching how stakeholders can determine the structure and operation of the accountability system will be useful to community colleges and others.

Accountability and TQM (Total Quality Management)

In the Webster's Universal College Dictionary (1997) accountability is the noun derived from the more commonly used adjective 'accountable'. Accountability refers to the obligation to report some matter or occurrence as reasonable/just/well grounded or legally warranted. It has typically been associated with obligations and responsibilities of the management of the financial matters of a company or other organization. After the 1970's people extended the meaning of accountability to include all matters (financial and otherwise) that defined the ultimate success of a company or organization as it was realised that companies that seemed financially sound (having stable sales, good profits and a relatively stable corporate history) could very quickly be decimated by a new market entrant who provided greater quality in the product or service and greater customer satisfaction (Juran, 1993). Thus the Japanese economy was moved in 20 years from being a ridiculed reflection of a defeated nation that was known for shoddy exports to one that destroyed the American pre-eminence in the manufacture of automobiles, colour television sets, electronic goods and more recently in computer chips and high technology items. This world quality leadership has been achieved in the American national market and globally as the result of the strategic application of Total Quality Management (TQM) principles (Garvin, 1983; Juran, 1993; Drucker 1999).

Juran has further predicted that if the twentieth century is synonymous with increased productivity then the twenty-first will be marked by a penchant for quality. Garvin (1983) especially emphasised that "—high quality means pleasing customers, not just protecting them from annoyances" and that "Some customer preferences should be treated as absolute performance standards." The history of quality is really about attitudinal changes in the way society has viewed the world of work. Quality was not taught in Western management and business schools prior to the 1970's but the first formal steps in quality arose from the infamous time and motion studies of an American, F. W. Taylor in 1912. Shewart developed the methodology of work sampling but statisticians Joseph M. Juran and Walter Deming in the 1940's and 1950's provided the theoretical base for understanding quality issues and fathered and nurtured TQM in Japan and the Far East. Philip Crosby did the same on a more

limited scale in America with his “zero defects” program for the federal government which defined quality “as meeting the customers requirements the first time and every time”.

By the 1980’s, TQM and its several variants such as Continuous Quality Improvement (CQI) and Continuous Quality Management (CQM) were present in many enterprises besides manufacturing. Besides service industries like banking, hotels and hospitals, it was used by organizations as diverse as the U. S. Navy, small retail merchants and large church congregations (Long, 1995) and was a part of studies in organizational behaviour (Johns, 1996) and management theory. Six Sigma is a later TQM version that has yielded fantastic financial and other benefits at Motorola Corporation and General Electric Corporation by providing mathematical modelling of how to control process variation and secure better design specifications (McFadden, 1993).

The TQM management style stresses the leadership’s commitment of the full human and other resources of the enterprise to the fulfilment of the customer’s satisfaction and even the exceeding of the customer’s expectations. It is assumed that the major source of inefficiency and of difficulty in achieving goals is due to the organization of the work process and not employee behaviour. Since the institution must reflect its performance of its mission by more than the value of a single factor, the measuring of institutional performance outcomes has become a major activity (U.S. Department of Education, 1998; Cabinet Office, 2000). Performance cannot be improved if it is not being tracked.

Educators have stubbornly resisted the direct implementation of TQM principles by administrators but have used it more often as a guide, though still at a pedestrian pace compared to its impact in the rest of society. Despite the surplus of TQM courses being taught throughout the nation very few colleges or universities have implemented it in their own systems (Spanbauer, 1994). Comparatively few case studies of TQM applications in education are reported in the academic literature (Thor, 1994) but in 1995 Lockwood published a study involving a sample of deans “from the 52 colleges represented by the Continuous Quality Improvement Network.” Beering (2000), speaking as the president of Purdue University, said “Excellence 21 is a system-wide effort by Purdue to explore the

principles of continuous improvement and total-quality management. In considering the use of management techniques that found their first application in for-profit organizations, Purdue is moving carefully into territory that is largely unexplored by higher education. Universities are unique enterprises and we would error grievously if we tried to re-cast them in the mold of corporations. Yet we would be negligent if we assumed we had nothing to learn from other successful organizations."

Application of quality in education has been given responsibility for success at all levels from rural high schools to departments in research universities. The big mental hurdle of accepting students as customers has been overcome by Deming's 1993 clarification that customer expectations are often caused by the provider. Thus the truly successful provider innovates, predicts and is in advance of the needs of the customers since the system of customer-supplier relationships is more interactive than customer driven. Such describes the preferred education situation. TQM was the vehicle for bringing greater success to the corporate sector, government and charitable organizations but has had little sustained impact in education. (Parker, and Slaughter, 1994; Dennis, 1995).

Sims (1992) traces a concerted demand for more formal accountability systems in higher education from the early 1650's (when the first universities and colleges appeared in the USA) up to the current stance of the regional accreditation agencies for colleges and universities as well as those professional bodies that accredit specific programmes and work experiences. William Bennett, later to be Secretary for Education, in 1985 warned that public colleges and universities should "state their goals, measure their success in meeting those goals, and make the results available to everyone" and "If institutions don't assess their own performance, others --- either state or commercial outfits --- will most likely do it."(Sims 1992). Yet progress in instituting accountability in higher education has been stoutly opposed. (National Association of College and University Business Officers Benchmark Program, 1998-1999; Parker and Slaughter, 1994). Though formal accountability systems exist at the tertiary level to some degree in about half of the states throughout the nation, they are most explicit in K-12 education (Wellman, J. V., 2001).

Calls for accountability, i.e. quality in education

Accountability in education first reached a groundswell in the 1970's with emphasis on fiscal responsibility of the administrators and productivity in the classroom (Kruger, 1970). This concerned inadequate returns on the inputs being made by the taxpayers. The call was for more graduates or at least more students who were successfully achieving the set standards of the system. Failure was blamed on the student since the ability to learn was considered to be some inherent, individual property and thus any lack of success would be due in great part to the student's attributes. Accountability was focused on increasing the returns for the money and resources allocated. Teachers declared themselves to be blameless and unable to shoulder any accountability until they enjoyed even greater control over their work. They insisted, via negotiations by the National Education Association (NEA), in fighting all attempts at any national accountability system of national testing or evaluation of students. Teachers declared they would need to have a major share in governance of school systems before they could even consider that the performance of students would in any part contribute to any sense of accountability on their part (Kruger, 1972; Polley et al, 1972).

But in the 1970's two events collaborated in a re-defining of accountability in education. One was a new observation and belief that all students could learn and that the capacity to be successful was very much in the hands of the teacher and the system and not wholly the responsibility of the student. The second event was the total 'open door' approach of the community college and its rapid expansion. This was an opportunity for a true expression of the democratic ideal and American truism of access for all to the good things of life. These good things began with a college education and the community college was 'the people's college' or 'democracy's college' (Roueche et al, 1972). The new paradigm was that all students could learn and accountability should therefore include parties other than the student as well. The system would be held accountable for equity and access issues. The administrators and educators would be held accountable for the provision of learning (Lessinger, 1970; Moore, Jr., 1971). In the community colleges persons, who were sometimes previously written off by educators at earlier stages of their lives, achieved notable success under appropriate learning conditions. The parallels of K-12 public education, like non-selectivity of students and the variable success within the system such

that some schools and districts accomplish highly superior results, are clear. Accountability would no longer centre on the student only. It would involve the teacher, the administrator, the school board and others. A myriad of candidate accountability schemes quickly appeared and often as rapidly disappeared. Holding the system and administrators and educators accountable was much more complex than holding students accountable and demanded much more circumspection. This is readily apparent in relevant articles in "Accountability in American education" (edited by Sciara and Jantz, 1972). America resisted real accountability by showing a lack of standards and meaningful action to support the words.

The stunning indictments of the education system in "A nation at risk" (1983) galvanized political and popular opinion. By the passage of the School Improvement Act in 1987, the various states moved voluntarily to share comparable data on education (Vinovskis, 1998). State governors met the outcry for school accountability by adopting eight National Education Goals with a yearly focus for national progress between 1990 and 1997 (National Education Goals Panel 1999; Council of Chief State School Officers 1999). The public absolutely refused to believe that education could not be held accountable and evaluated to give the same benefits in productivity and efficiency as had been seen in business and elsewhere (Finn, Jr. 1999; Immewahr, 2000; National Partnership for Reinventing Government 1999; Manno, 1998; Mitchell, 1996). The taxpayer is the major funder of K-12 education through local taxation and thus has a particularly strong influence on the school's response to the Accountability Mandate that had been required of the politicians (Morris, 1972; Cunningham, 1972). Educators in the tertiary institutions bemoaned the quality of students they were receiving and blamed the extra time they had to spend in remedial activities for the spiralling costs in higher education. This impacted the public funding purse and also affected the personal wallet because of increasing tuition and fees (American Council on Education, 1998). Meanwhile, employers complained of the low quality of high school and sometimes college graduates and the lack of a quality workforce was an impediment to improvement in local economies (Darling-Hammond and Ascher, 1991; Immewahr , 2000).

A new accountability that focused on the quality of education by whole school reform was intended to "help all students, even those on the margins, succeed in school." (Mc.Chesney, 1998). It would address the quality of the school experience and the learning opportunity of the system rather than just seek increases in numbers on a checklist. Fuhrman (1999) reported that, "As part of standards-based reform, states and districts are designing new approaches to holding schools and districts accountable for discharging their missions." and "Most are taking the next step which is to use achievement of the standards as a basis for accountability." Several organizations such as the Thomas B. Fordham Foundation, the Center for Education Reform, the Annenberg Institute for School Reform and others have acted as powerful resource and lobby groups to buttress the drive begun in the 1986 Governors Summit at Charlottesville. Nationally, the federal government has assisted private and quasi-government agencies to set and review defensible standards in the main academic curricular areas of science, history, civics, geography and the arts. Additionally, the standards in the core competencies of reading, mathematics and writing are set at regional levels. The Council of Chief State School Officers and the Education Commission of the States encourage and promote exchanges and collaborative efforts on standards and assessments (Allen, 1994; Barton, 2001). The standards-based accountability schemes placed less emphasis on compliance with regulations to gain accreditation or certification. Previously state officers would track compliance by self-reporting by schools and districts of pupil to teacher ratios or site visits to confirm curriculum adoption by inspection of school board minutes. The new criteria required the monitoring of outcomes data such as school graduation rates; attendance and drop-out frequencies and other non-traditional indicators. Accountability is not only based on the three R's but also the level of performance and the level of improvement in the other academic areas as well as the new indicators (Fuhrman, 1999; Goff, 2000; Sirotnik and Kimball, 1999).

Consequently, several accountability schemes (Finn, Jr. and Petrilli 1999; Doyle 1996; Fuhrman 1999) have standards that all schools must meet. Low performing schools are penalized for performances that are consistently poor or which are not progressing at an acceptable rate (Lynd 1996; Chronicle of Higher Education March 2, 2001; National Center for Education Statistics, 2000; State of State Standards, 2000; Florida Department of

Education, 2000). In 1998 the American Productivity and Quality Center (APQC), the National Alliance of Business and the Council of Great City Schools formed a collaboration for a first-of-its-kind study "to identify and examine innovations, best practices and key trends in the area of accountability systems"(American Productivity and Quality Center, 2000). The APQC now carries a regular section devoted to K-16 education (four years beyond the normal K-12) and it is therefore promoting a seamless system to be devoted to quality performance through accountability from K-12 through to the community college and even the full undergraduate degree institution. The APQC therefore provides a forum for discussing and comparing accountability issues in business alongside those in education (American Productivity and Quality Center, 2001).

In summary, educators initially did not accept the direct application of TQM but there were strong proponents in the business leaders and government officers who either participated in or influenced the boardroom dealings of the learning institutions. Responding to alarms from many constituencies about the costs and low productivity of education the politicians exerted pressure through the State Governors Conference and also areas like the Council of Chief State School Officers (CCSSO) to impose an Accountability Mandate on K-12 education. Varying responses followed (Finn, Jr., 1999; RAND Research Brief, 1999; State Education Accountability Reports and Indicator Reports, 1999). The work of the APQC exemplifies a collaboration of business and educational experiences to drive the accountability agenda in education. A constant of these new schemes is the focus on continuous improvement; on incremental gains rather than fantastic swings; and on prolonged and sustained enhancement as indicators of success in achieving the quality of performance desired (National Educational Goals Panel, 1997; Center for Education Reform Action Paper, 2000).

The purposes of accountability systems

To develop a necessary accountability system the politicians either called together panels of experts or designated some branch of the state's education authority (perhaps like a co-ordinating board) to lay down guidelines and procedures. Educators avoided being directly involved (California Assessment Institute, 2001). These facts are important because they affected whether accountability would be primarily formative, summative or a mishmash. In

a summative system the focus is on simply collecting and collating information to meet certain designated values. This is reminiscent of the number crunching that marked the old simplistic fiscal accountability. The information is simply compared against some pre-set values. Certain consequences follow from satisfactory values and other consequences from non-satisfactory values. This type of accountability is seen when using an accrediting agency checklist. It is easy to administer and understand and meets the original political ends that drove accountability.

Formative accountability is a much richer procedure to address the higher objectives of education. The information feedback is used to push improvement. The objective is to relate the outcome value to all or any other circumstances as being causal or an effect or there is no relationship. A chain of causality will eventually be established such that the factors that are accountable for the observed outcome values will be identified and understood. A model can then be developed that will allow predictions or extrapolations such that process controls are understood rather than just followed by taking measurements. For formative application it is important to have as much information as possible both in terms of historical data (so as to look for trends) as well as ancillary data so as to search for unexpected correlations and interactions. This will necessarily involve a more complex system as it is more than just collecting numbers for check off against set values. Because it is intended to explain what is happening and to give feedback, a formative accountability system is especially useful, however, to educators and administrators; but it is also useful to all who have the courage to examine it in an open and fair-minded fashion. Formative evaluation may suggest relationships that direct courses of action that are unpopular. A formative accountability system is meant to promote direct accountability (Sims, 1992).

In K-12 in the USA, accountability has been summative primarily. In the 1980's the National Council of Teachers of Mathematics (NCTM) attempted to have national standards by using standards-based reform to raise mathematics performance. The NEGP promotion of Goals 2000 was a systematic response to the Accountability Mandate from the state governors for content standards in the subjects; performance standards by the education systems and comparable or aligned assessments for all students (Barton, 2001; National

Education Goals Panel, 1999c). Since 1990 state-by-state data have been collected and collated from the National Assessment of Educational Progress (NAEP) of the Educational Testing Service (ETS). NEGP officers systematically analysed and regularly reported this data in a format that allowed ready comparison between states and which showed trends by noting the historical data (NEGP Monthly, 2001; National Education Goals Panel, 1999a). A case was being made for a national testing system (Davey, 1992; National Education Goals Panel, 1999b) as an accountability system but there was still concern that accountability should be controlled locally (Bowers, 1989). Whether to be nationally organized or at the state or district level, the major thrust for accountability systems was by testing of students to see what they had learned. This was summative as results were screened against set targets or "National Education Goals".

Unusually, the National Partnership for Excellence and Accountability in Teaching made the improvement of teaching its core mission and led research "to guarantee that standards and assessments really measure teachers' effectiveness with students and that new assessments are fair and appropriate." (Hawley, and Valli, 1998). A formative accountability system was implied and the presence on the Policy Board of representatives from the American Federation of Teachers, the American Association of Colleges for Teacher Education, the American Association of School Administrators and others suggested a chance for real collaboration and trust in time. Jennings (2001), speaking as the Director of the influential Center on Education Policy, in an open letter to President G.W. Bush and the Congress on the first of the Core Principles for an Improved Federal Role in Education, implored that "The federal government should continue to encourage high academic standards, but should also demand meaningful accountability from the states for increased student achievement and accept national responsibility to help in the proper use of tests."

Overall, some saw test results as the primary accountability whilst others referred to high standards with a separate reference to 'meaningful accountability' that was distinct from testing. Accountability was a response to a nationally recognized problem of declining standards and a national reference or comparability system was needed (and agreed and initiated by the Governors' 1986 Summit). In a decentralized system that had local tests of

highly variable and sometimes highly suspect quality, NAEP data were an early beacon in the confusing murkiness between no local system or low quality systems and the need for comparability of outcomes, i.e. comparable accountability systems.

Some state accountability systems.

Darling –Hammond and Ascher (1991) identified Political Accountability (Legislators and school board members were elected); Legal Accountability (The Courts and legislatures enacted policies and laws, and responded to civilian complaints about education); Bureaucratic Accountability (The rules, guidelines and procedures that were the minimum that the school system met to satisfy accreditation or funding, etc.); Professional Accountability (Educators possess various qualifications and pass certifying exams and should, as true professionals, hold themselves accountable by their own self-regulating bodies as other professionals do.); and Market Accountability (Given full information and full access and equity, individuals/customers would choose the school or system that best met their needs). Bureaucratic accountability systems have been the most prevalent. Several writers (Finn, Jr. , 1999; Manno, 1998; Stapleman, 2000) have called for leadership from the federal and state governments (Political Accountability) to address the proliferating variety of systems that have appeared. Clearly, however, the new results-based accountability systems offered opportunities and challenges in many areas (Horsch, 1996; Center for Education Reform Action Paper, 2000).

Such systems implied formative accountability. They supplied copious information and called for institutions to be responsible for initiating the necessary action and following up on results. This went beyond the passive submission of facts and numbers in traditional input measures systems. The serious commitment of the states was seen in action such as the California Public School Accountability Act of 1999. By 1997 there was sufficient history that a review of standards-based reform in 9 states (New Jersey, California, Florida, Georgia, Kentucky, Minnesota, South Carolina, Connecticut and Texas) was in the opening section of a comprehensive report to the NEGP on “Implementing Academic Standards”. Texas and California were expected in this list with their large populations and huge education systems with commensurate challenges. The spread of activity on accountability was seen in the

appearance of states like Georgia and Minnesota. Test-based systems could have problems of teaching to the test without real learning (Stapleman, 2000) and Kentucky was an early mover in addressing this challenge to standards-based parts of the accountability scheme. (RAND, 1999). Thus, despite fears of an over-emphasis on standards-based testing (Olson, 2001), some states (Florida and Kentucky) followed a more pervasive system that, besides student accountability, also directed accountability attention to educators, principals, school board members and aspects of the physical structure and organization of the school (Sandham, J. L., 2001; Florida Department of Education, 2000). It is noteworthy that these more progressive states also provided the requisite professional development to assist persons to respond appropriately to the accountability demands.

A full report on state systems from Finn, Jr. and Petrilli (2000) and also from the annual reports of the CCSSO State Education Center (1998; 1999) gave much cause for acclaim. All 50 states had operating systems with state standards and some published accountability responses for systems that did not meet satisfactory performance levels. Texas was included in a select group of 5 states as excellent role models of having strong accountability systems to back up solid standards for the tests. Movement to develop accountability systems for special circumstances within a state produced efforts in 2000 by the Great City Schools to promote a trial NAEP assessment for large urban school districts with collaboration by the American Federation of Teachers, a union that had shown very little co-operation in the accountability movement. Details of individual systems were available usually from the offices responsible for education oversight in the individual state and through CCSSO on-line publications. The NEGP produced substantial support materials for any interested parties. Much still remained to be done before all states were level with Texas and others (Finn et al, 1998;Viadero, 2001) but a Whitehouse Brief on Educational Progress 1992-2000, (issued late 2000) showed that in 1996 only 14 states had standards in core subjects and 4 years later 49 states had such standards. Individual states had come a long way with some travelling faster than others.

The Texas education system

In 2001 Texas had 35 public universities and associated academic centres, 3 public state colleges, 50 public community colleges, one public technical college system with 4 campuses, 7 health related institutions, 37 independent 4-year colleges and universities, 2 private junior colleges and 1 private medical school. These institutions enrolled 990,403 students in Fall 2000 with 90% being in the public system (Rylander, 2000; Texas Higher Education Co-ordinating Board, 2000). The public universities and the community colleges shared the pool about equally. Texas had 2 public institutions ranked in the top 50 nationally with 11 of its Ph. D. programs being ranked in the top 10 nationally. In 2001 the University of Texas at Austin was the biggest university in the country with over 50,000 students but the University of Texas System had plans to create 'centers of excellence' in existing universities to upgrade them so that at least one more Tier 1 institution was produced for each of the next three decades (The University of Texas System, 2000). Similar aggressive strategy was to dramatically enhance research capacity in the 5 largest University of Texas System health components.

In 2001, Texas's population of about 19 million was expected to increase by a further 14 million by 2030 with most of these individuals (about 90%) being non-Anglo. Texas had 3 of the 10 largest cities in the nation. The very buoyant economy had moved from an agricultural and mineral extraction base to one dependent on high technology ventures associated with computer and various information-related specializations. Higher education was funded through income from oil producing lands (the Permanent University Fund) as well as revenue allocations by the Texas Legislature. This gave the Texas Legislature great clout in terms of accountability. For higher education, the state only recognized institutions that had been properly accredited by The Southern Association of Colleges and Schools (SACS) or an equivalent. The Texas Higher Education Co-ordinating Board (THECB) was empowered by the Legislature to approve curricula, collect and collate information on accountability (often called institutional effectiveness measures). The 18 THECB members were selected state wide by the Governor but could not be a trustee of any college board or be employed in education. Community colleges had locally appointed, individual boards

whilst the huge systems operated by the University of Texas, and Texas A and M had Boards of Regents appointed by the Governor (Texas Higher Education Co-ordinating Board, 2001).

In summary, the Legislature held all institutions to a standard of accountability through the work of the THECB, through the SACS accreditation exercise and through certification procedures and licensure protocols for various specific programmes and careers. Institutions were accountable to the extent of their dependence on public funding. Thus community colleges were very dependent and other institutions were much less so. There was no Master Plan for Education in the sense of the California Plan and indeed the "Agency Strategic Plan for the Fiscal Years 2001-2005" which was published by the THECB emphasised a minimum of government involvement. Data on Texas Public Universities and also a Statewide Factbook for Texas Public Community and Technical Colleges were published annually and also available on-line but were not generally accessed or used much by the public according to THECB officials.

This higher education system was built on the K-12 system. Texas K-12 public education in 2001 involved 3.95 million students on 7,228 campuses in 1,103 school districts in 20 regions. The 10 largest districts had at least 50,000 students each (22.2% of all public school students) while the smallest districts had less than 500 students each and although they were more than a third of the districts they enrolled only 2.5% of the students. The system was typified by lots of districts with small numbers of students on three campuses at most --- one each for elementary school, middle school and high school. The range is illustrated by the fact that the Houston Independent School District in 1999 served 210,000 students on 296 campuses while Divide Independent School District in South-Central Texas had only 19 students. Typically, major urban areas had 80+% minority student body (where minority was defined as any of African American, Native American, Hispanic, or Asian or Pacific Islander ethnic groups) but the trend was reversed in most rural areas except for the areas in the south and west, which bordered Mexico and where there could be very high Hispanic populations. Hispanic students were the largest overall minority at about 38% and African Americans were about 14%.

All K-12 education was funded locally by local property taxes and there was no state income tax. Recent governors had cut school property taxes deeply and raised state supports to equalize spending in local school districts. In 1997 public K-12 funding had been increased by more than 50 % of the budget increase over the previous biennium with a 37% rise in the state expenditure per student. The State Board of Education acted as the policy making and strategic arm of the Legislature for K-12 public education while the Texas Education Agency was the body that accredited school districts, dispersed state and federal funds, administered state-wide student assessment and other accountability measures, oversaw curricula and engaged in a host of other supervisory activities.

Texas's reputation for the rigour and success of its accountability system (National Educational Goals Report 1999; National Education Goals Panel 10th. Anniversary Report 1999) was national. Texas had had a long history of uneven development educationally. The mixed agricultural/mineral exploitation economy was ably supported by having basic education levels and the old pattern of access by White males to the best facilities. A system of fewer opportunities with less concentration on education of females and non-Whites was allowed to hold over from Jim Crow segregationist patterns (Lynd, 1996; Scheurich, 2000). Dependency on local funding with the continued absence of non-Whites in the highest socio-economic levels and better school districts meant there was further inequity of access to education. Despite there being over 50 education goals for public education in 1994 more than 20% of Texas schoolchildren could not read at their designated grade level, drop-out rates were high and social promotion from one form to the next (regardless of the lack of performance) was the norm. Governor G. W. Bush made improving Texas's education his number one priority on taking office in1994. He advocated fundamental system reform of education with local control and accountability as key tenets. He therefore changed the focus of TEA to carry out what he saw as the state's responsibility to set high standards for the education system and to support that system and hold the schools accountable both fiscally and for student achievement. This latter focus was a change that was long overdue. Clear goals were set for excellence in Mathematics, Science, English and Social Studies and local parents, educators, administrators and school boards were to have local control of the school.

Accountability systems in 2-year colleges

Over 90% of two-year colleges (public community and technical colleges) were funded in the main by federal, state and local means (Vaughn, 2000). Sharp separation between the K-12, 2-year college and 4-year college/university is a distinctive feature of the American scene but there are some similarities between K-12 and the 2-year colleges that are useful considerations. Founding personnel in the 2-year colleges often came from an experience in the secondary school portion of K-12 and many of these individuals have persisted until the present as faculty and administrators (Cohen and Brawer, 1996; Personal notes 2000, 2001). The tradition of union involvement and many of the negotiation areas found in much of the 2-year arena are a carryover from K-12. Senior 4-year institutions had little or no union involvement in professional matters and relied on a very effective and active collegially and discipline based self-regulatory strategy (Personal notes of Intersession 2001). Tertiary educational systems as a whole had not been brought under the strictest gaze of the accountability mandate as had K-12 education by the dawn of the 21st century. Further discussion is restricted to the 2-year college situation.

Despite vociferous urgings by business persons associated with the colleges and such seminal works in the field as "Blind man on a freeway: the community college administrator" by Moore, Jr. (1971) and "Embracing the tiger: The effectiveness debate and the community college" by Roueche, Johnson, Roueche and associates (1997) there has been only fitful movement to embrace accountability. In the first book the administrator is challenged to hold himself accountable for really being able to deliver a quality service so that the 'people's college' or 'democracy's college' would indeed be accountable to the people. In the second book one is exposed to examples of institutions where insights and strategies from the accountability movement have been successfully employed. Most regional bodies responsible for college accreditation and those which carry out accreditation of programmes and courses now have moved to outcomes-generated assessment (Sims, 1992; Texas Higher Education Co-ordinating Board, 2001; Santa Monica College, 1999; Missouri Department of Higher Education, 2000; U. S. Department of Education, 1999). Yet in 1971 Roueche, Baker and Brownlee opened a discussion entitled "Accountability in the two-year college" with the following sentence. "Accountability is becoming an increasingly popular and controversial

concept among educators." Despite their very cogent arguments about the need and consequences of embracing it or holding off from the clear exercise of accountability, the major driving force had come from state legislatures like Texas, North Carolina and California (The California Assessment Institute, 2001; Ewell, 2001). Texas had very tight collaboration of the wishes of the legislature and the response of the community college since funding and approval to initiate programmes, etc. were tied to the degree to which the colleges satisfied the Texas Higher Education Co-ordinating Board. North Carolina went further and the community colleges compete directly for funding based on accountability criteria (performance-based funding). California had a more complex situation because of the involvement of shared governance whereby faculty and students actually shared directly in the governance of the institution. Here the accountability mandate tended to show the least degree of development. In their book "Accountability in American education" Sciara and Kantz devoted a chapter to "Applied Accountability" and through some 70 pages various authors described a situation as regards accountability in the 2-year college that is only marginally different from today's scenario of thirty years later.

Some strident indication of change is on the horizon, however. Florida is accredited by SACS (like Texas) and was considered to have a fairly progressive accountability system. Some have countered that it was not sufficiently responsive to the state legislature and the whole regulatory structure was in the process of being revamped in 2001 (Personal communication from David Armstrong, 2001; Notes from Community College Leadership Program 2000, Notes from 20001 Intersession). Using the momentum and experience gained from the K-12 accountability mandate politicians and the public no longer recognized the autonomy that these colleges once had. Without accountability public support will be lost and so accountability systems were themselves being reviewed and assessed (Wellman, 2001;Ewell, 2001; California Assessment Institute, 2001; Callan, Doyle and Finney, 2001). In 1999 the American Association of Community Colleges developed a report called "Core Indicators of Effectiveness". Perhaps this will be the bold and definitive start for which many had been calling.

Some insight can also be gleaned by considering accountability in overseas education venues. In America publicly funded K-12 education is the norm and the 2-year college is the commonest first access to public higher education. The private sector was always an addition to public education. Public education is a state responsibility according to the Constitution although as of the 1950's there has been increasing federal involvement. The system is also highly decentralized with any associations of schools or colleges, etc being on a voluntary basis. The state licenses/certifies educational establishments to operate and usually uses one of the regional accrediting agencies or enforces criteria by a state department of education or similar. The U. S. Department of Education exists as the first among equals and has its greatest role in acting as a contact and reference point at the national level for international relations in the education sector (Vinovskis, 1998).

Contrarily, most other nations have a centralized system of education that is controlled tightly by the central government with some variation in the degree of individuality. Public education was not always so predominant and private education has played a major role in countries such as Canada, Japan, Brazil and South Africa (OECD/CER! 2000). Here the standards and criteria are still set by central government for both the private and public sectors and all levels are under central control. Access to education has expanded in other countries to approach the rates seen in the USA. Virtually no other countries had open door access to higher education as existed via the American 2-year college but almost all these countries had uniform graduation criteria to indicate successful completion of education at the secondary level. Facing the spectre of rising costs at home and abroad, this meant that education demanded more from government at the same time that these countries were improving their quality of life to be closer to that of the USA. The USA and Europe had to support fewer overseas students and in some cases there was a limit put on foreign entry to some programmes because of increasing demand by their own students. Thus due to rising costs, accountability also became a major issue in many overseas jurisdictions after the 1960's (Dill, 1999; Cuttance and Stokes, 2000).

Overseas, the fact of centralized control had most often meant that once the case was made for accountability the debate had been about what form it should take and not whether or not

it applied in each particular case. There had been far less of the instance of pleading 'special exemption'. Much of the attitude as regarded accountability was proudly acclaimed as originating in the quality movement (TQM and its offshoots). These were seen as "means for restoring and strengthening the internal web of accountability by which colleges and universities have traditionally assured the quality of their teaching and learning" in ways that would be similar to the means by which the guilds and craftsmen's associations had maintained quality standards of work. With this tradition of collegial review and self-assessment but limited or no external or professional accreditation, the system of academic audits was developed in the United Kingdom. This quality assurance mechanism focused on institutional academic standards. All results were public information and published and ranking tables could be generated. Asian Commonwealth ties, Hong Kong, New Zealand and Australia were soon experimenting with it. In Europe it was amended in Scandinavia, the Netherlands and a few other countries (Dill, 2000; Trombley, 2000; Department of Further Education in England, 2001). School performance was published in 'league tables' that were available for inspection by parents and students (British Broadcasting Corporation News, 1998) in the United Kingdom. In Australia there was much active research about how parents were given accountability information for their children and what use they made of it (Cuttance, P. and Stokes, S.A. 2000). In Canada, the use of performance indicators to decide funding at the post secondary level had shifted the requirement for demonstrating quality in education to the individual institution (Association of Colleges of Applied Arts and Technology of Ontario 1996, 1996; Barnetson, 1997). For statistical insight into student achievement and educational systems on an international basis "The International Indicators: A time series perspective, 1985-1995 " was particularly informative (National Center for Education Statistics, 2000).

Rationale for a study

The Texas accountability model has had a number of critics (Klein et al, 2001; Yardley, 2000; Haney, 2000; Clark, 1998;) as well as persons who tout its success (Skrla et al 2001; Sclafani, 2001; Jacobs, 2001; Finn and Petrilli, 2000; Strong, 1999; Johnson, 1999). Its comparative longevity over other accountability systems and its relative constancy and consistency of structure have made it unusual and admired among the plethora of constantly

changing systems (Scheurich, 2001). In common with most other systems, there has been an absence of input by various persons during its design and development and there is a continued absence of the voices of various stakeholders in the education system in the ongoing debate about the success of the accountability system (Rothstein 2001b; 2001c). There has been vociferous and heated clamour from political and academic sources about the performance of the accountability model and the consequences of its application (Bello, 2000a; 2000b; Sanger, 2001; Wilgoren, 2001).

A search of the literature did not reveal any previous material that examined the public's opinion about the performance or otherwise of the education accountability schemes. Yet these schemes have been developed supposedly to serve the public interest. There is, however, much literature on the quality movement and quality tools and how all this has strengthened accountability from the customer perspective and led to improvements in product service and quality as well as advances in productivity, efficiency and effectiveness for the manufacturer as well as the service provider (Hauser and Clausing, 1988; Juran, 1993). Such customer-centric accountability has led to a win-win perspective. Therefore, this project sought to develop and test a process to allow various stakeholders to express their opinions on the education accountability schemes from a customer-centred view. This is appropriate since public education is for the public benefit and therefore public satisfaction with those means that are used to assure its quality would be important.

Accountability systems can drive an organization to improvement and the attainment of the highest quality in the production off goods and services (Juran, 1993; Menezes, 1991). The history of the management of quality has been about holding people, processes and systems accountable (Class notes; 2001 b). It began with the quantitative methodology of Taylor and then Shewart at the start of the twentieth century. The infamous stopwatch type time-and-motion studies of Taylor's 'scientific' research of work processes were focused on producing quality programming to raise manufacturing productivity. After success in manufacturing, it was the genius of Juran, Deming, Crosby and other American quality pioneers to show that quality could be secured in administration and service areas as well by the same principles of accountability. Garvin (1995) traced how quality measures evolved beyond the 'purely

statistical controls on quality'. The greatest quality and productivity gains were linked to maximal long-term satisfaction of the customer (Lahiri, 1999; Baldrige National Quality Program, 2001). In order to better please customers it became very important to get information directly from the customer. Some of this was quantitative data but far more of it was qualitative data.

A bedrock principle in producing and managing quality performance is for the enterprise to be customer focused and to have the accountability system centered on that same objective. All activities are geared to meeting the customer's immediate expectations and then exceeding them to produce 'extraordinary quality'. With inadequate research on customer requirements a gap develops between the service delivered and the customer expectations because of the disconnection (Berry et al, 1994; Menezes and Sorbin, 1993). This customer centric research requires both quantitative and qualitative methodologies. It reveals solutions that can generate the full customer satisfaction that is both attainable and valuable (Hart, 1988; Hauser and Clausing, 1988).

The research gap

The literature held no direct research on the evaluation of accountability systems in K-12 American public education from the viewpoints of students and their parents, or of employers and business persons, or of higher level educators. This is not unexpected because of the long tradition of the fragmented nature of American education, which tended to make accountability a local issue and accountability in education was not a high political priority for much of the time --- especially in the Southern states (Black and Black, 1987). The various states, historically, did not collaborate or share information about education except on an ad hoc and occasional and totally voluntary basis (Vinovskis, 1998). Indeed, even in the same state, K-12 education and higher education were often two non-communicating entities. This had begun to change by the 1970's such that as demands for state accountability in education arose, states like Florida mandated that there would be collaboration between the state education agencies that allowed easy articulation of students moving between educational levels (Council of Chief State School Officers, 1998).

Any study of accountability in education starts from a history of it being very localized (Black, E. and Black, M, 1987). Responsibility was operated at the level of the school board or school district managers and the state and federal government had successively less influence. The system was in reality highly undemocratic, as although members ran for elected office, they never espoused equality of educational access or quality as a feature of local politics. This existed until the advent of the Civil Rights Movement when quality and access issues did take centre stage (Hacker, 1995; Fuchs, 1990). The first impediment to a general exercising of strict accountability was that the average American was uncertain about how to hold the education system accountable. There was no such expectation from history or politics but the situation with regard to accountability practices has been changing. Much more research is being conducted at present in the arena of equity, access and achievement in education (Scheurich et al, 2000, 2001; Fowler, 2000; Rose and Gallup, 2000; 1998).

A second impediment arose because the consumer model of accountability came from the business world. Total Quality Management (TQM) in the American landscape was tied to rapid gains in a declining economy in the 1980's, but it faltered badly in education. When TQM was attempted to similarly transform the educational system unionised teachers and others, who were uneasy about the idea of students being treated/considered as customers, vigorously opposed it (Parker and Slaughter, 1994; Sciara, 1972). In America education was the great egalitarian leveller to upward mobility and success for anyone. It was too sacrosanct to be seen as another commodity in the dictates of business theory. Thus, education generally has missed TQM developments that have emphasised the role of the customer's expectations and customer satisfaction (Spanbauer, 1994). The various parties in the K-12 public education system seldom have or understand organized processes for examining the quality and performance of the system. There have been successes attributed to TQM and CQI (Roueche et al, 1997; Scheurich et al, 2000) but generally those responsible for the education system have discouraged or ignored the use of such management and decision-making tools.

The major tool that the average stakeholder had was the system of accountability ratings (the Texas model) and anecdotal evidence. Consensus on the exceedingly good or atrociously

bad schools in any area was easily supported by the success stories or horror recriminations that built up over time. Such methods were clearly crude and very subjective. By contrast, under the political mandates of the governors of several states, K-12 educational accountability was to be a very objective assessment of performance as measured against various indicators set by state agencies. These indicators were produced with little educator input (due to the union's general refusal to participate) and even less by the public. There was no consumer concept of education (Texas Education Agency, 2000). There was no tradition of broad collaboration about education.

Conclusion

Accountability had long been talked about but only slight progress had occurred by 2001 in the USA in setting up formal accountability systems in higher education. More progress had been made at the K-12 public school level than at either the 2-year college or the senior institutions. The senior institutions had maintained their tradition of autonomy as fostered by the Constitutional relegation of education as the responsibility of the individual states with only a very limited federal role. The 2-year colleges had resisted formal state accountability but because of their dependence for funding on local and state sources they would have to eventually yield. The situation varied across the country. Performance funding and accreditation links to accountability were being used to push compliance that tended to be grudging at best.

By contrast, originating with a meeting of State Governors in 1986, the National Education Goals Panel formed in 1990 and set guidelines for an Accountability Mandate. Goals with timelines were set for all states and the results and standards were to be compatible and comparable. This yielded a number of accountability systems that were outcomes-standards based with centralized testing of students being a focal and recurrent facet but not a sole criterion. Under this type of scheme the school as a system, the administrators and educators, as well as the students were all evaluated. Accountability systems had been established across the country in a comparatively short time of about a decade, but there was variation in the rigour of the standards and the vigour with which the accountability system operated.

The Texas accountability system was hailed as an effective model because it was both rigorous and vigorous.

The position is advanced that the accountability situation in K-12 is a premonition of what will transpire in the community college and therefore it is worth careful watching. The community college has however an even closer relationship with business and industry than does the K-12 system and some persons, particularly those with experience in industry and or teaching relevant management programmes are already very familiar with the quality movement. Atkinson-Grosjean, and Grosjean (2000) noted the convergence of performance-based assessment models for higher education on the international stage. Everyone has experience of K-12 public education and the community college is the 'open door college' that is to be freely accessed by everyone. The public did not raise concerns about access to either K-12 or the community college. Questions were raised about the quality of the systems, however.

The project therefore pursues how the public might become involved in the entire process of accountability. The quality movement suggested that the customer or the consumer of the service had a critical role to play in defining quality (Juran, 1995; Garvin, 1987; Meister, 1999). The federal government had led the effort to improve customer satisfaction by setting Customer Satisfaction Standards for all dealings with government. Specifically, certain funds such as Carl Perkins Grants were tied to appropriate accountability systems, which included a measure of indicators of satisfaction by the students and employment community. As early as 1970 Lessinger had actively promoted more public involvement in holding schools accountable. Support to enable the public to access the tools and become comfortable with their new active role in the accountability movement was developing rapidly (American Quality and Productivity Center, 1999, 2000, 2001; Center for Education Reform 2001; Business Roundtable, 1996, 1998, 2000; Standard and Poor's School Evaluation Service, 2001).

Chapter 3 — Methodology

Outline

Case study methodology was used to explore input from various stakeholder assessments of an accountability model. A self-administered questionnaire survey was hand delivered to two groups (educators and parents) and generated about 150 responses about (i) satisfaction levels with an outcomes-based accountability model; and (ii) specific stakeholder criteria that would be considered central to allowing the stakeholder to evaluate the success of the accountability system and to assist with the improvement of the accountability system. The information from the questionnaire sets was collated to generate a thematic overview of the responses about the assessment of accountability.

Parents are usually considered the least powerful of any of the stakeholders involved in the education system but yet they make the most contribution (by sending their children to school, by politically supporting the system, and as taxpayers). The education system has long convinced parents and most other stakeholders to have a hands-off approach (Sims, 1992; Sciara, 1972) and only recently have politicians become actively engaged on an extended basis (Fowler,F. 2000; Vinovskis, 1998). Some researchers have been calling consistently for educators to step forward and embrace the philosophy of accountability and effectiveness (Roueche et al, 2001; Roueche et al, 1997; California Higher Education Policy Center, 1996; Moore, 1971; Lessinger, 1970;). Parents often have had their decisions or opinions formed for them or heavily impacted by the experts of the education system ---the teachers and administrators (author's personal experience as a student advisor). Therefore parents have generally not had much direct input or a prominent role in developing the accountability models (Texas Education Agency, 2000; Fuhrman, 1999; Doyle, 1996; Finn, Jr. and Petrilli, 1999. It would be interesting to see how the views of parents, when explored in some depth, relate to those of the other stakeholders.

Choosing a research tradition

A common misconception is that research must be either qualitative or quantitative. Both research types involve systematic and logical thinking and then methodical investigations that are intended to increase knowledge about the topic (Hussey and Hussey, 1997).

Creswell says that the positivist (scientific) viewpoint, and its accompanying association with quantitative research methodologies, has a longer tradition than the naturalistic or post-positivist paradigm. This scientific method is commonly thought of as the most appropriate to give an unbiased answer to important questions (Slife and Williams, 1995). It is associated with quantitative data analysis since numbers is the language of positivism. Many social science phenomena do not follow positivistic restrictions as regards the observations.

Most state accountability systems for public K-12 education in 2001 were based on securing quantitative data (NEGP Monthly, 2001; Texas Higher Education Co-ordinating Board, 2001c; Chronicle of Higher Education, 2001). These came from students outcomes assessment based on the number of students securing certain minimal numbers of points in a certain set of subject areas (Finn, 1999) --- i.e. testing-based accountability. Some data were secured from other numeric performance indicators (Finn, Jr. and Petrilli, 2000). In the nearly global accountability mandates of the late twentieth century such accountability schemes replaced earlier forms of accountability, which had been based on numerical assessments of inputs (Sims, 1992). These had focused on the number of books in the library per student; the ratio of the number of higher-level degrees among staffs to the total number of students; and other such criteria. The major movements to assess performance and quality in education, i.e. to hold systems accountable, have relied on quantitative methodologies to establish the unvarnished and unbiased facts in a truly democratic manner (Slife and Williams, 1995). Yet school system accountability has a very strident political legacy (Dorm, 1998; Vinovskis, 1998; Bello, 2000a, 2000b). It is essential that any research should be as complete as possible and numbers alone often do not paint a complete picture, as noted by King (2001) in citing Patton (1990).

In summary, this research project explored the development of an assessment scheme for stakeholder requirements for K-12 public education by both quantitative and qualitative research methodologies. King (2001), Edwards (2000), and Hussey and Hussey (1997) have reported other theorists besides Patton as confirming the benefits of using such a combined methodology. The development of the study of management of quality by fostering

accountability and implementing accountability systems had its genesis in the business world. As the principles of customer satisfaction and the generation of quality in service delivery and administration have been investigated, the research has moved from being primarily quantitative to using both combined (Hart, 1988; Menezes and Serbin, 1993). These supported the use of a mixed research methodology here.

Research Methodology ----case study survey

The specific research methodology was a part of the phenomenological tradition with aspects of both an exploratory and an explanatory case study. It was exploratory as it was not an example in a series of such studies but was a single venture or exploration into the research area. It was explanatory in that it was intended to bring about understanding of another case (Texas) and to contribute further knowledge and understanding to theory in general. The methodology was identified by using Sutherland's Selection of Methodology Decision Tree according to Northcutt (2001). Such identification defined what tenets and strategies would be used for data collection and analysis but it was not allowed to constrain the progress of the project. According to recommendations from Hussey and Hussey (1997), the research must be guided by the principles of the genre to which it belonged but it must remain flexible enough to meet real world circumstances. The particular circumstances of this research showed such a trend.

The exploratory stage

The exploratory nature of the project examined how, in a separate case study, the various stakeholders who parallel the American situation, could show such stakeholders in public K-12 education how to evaluate the accountability of the education system. It described a different paradigm that was based on using greater accountability to the customers' or stakeholders' perspectives to drive improvements in the productivity and the operation of the system. The project was aimed at developing a system and instruments to secure input from the case study stakeholders on their perspectives of the accountability model.

The explanatory stage

The major tenets of K-12 accountability in America were set at the end of the twentieth century and progress towards the National Education Panel Goals has been slow compared to countries in Europe and Asia (National Education Goals Report, 1997; Dill; 1999; National Crosstalk, 1999). Such countries each had a unified and centralized educational system and adapting the best precepts from TQM and similar data management/decision strategies was much easier. American educators looked to them for the best paths forward (van Raemdonck, 2000; National Cross Talk, 1999).

Generally the American public's weak confidence in the accountability strategy for K-12 public education (Gifford, 2000; Patterson, 2000; Rose and Gallup, 1999) contrasted with their satisfaction and confidence in the public education system (Rose and Gallup, 1998, 1999, 2000). This was seen in the proliferation of alternatives to public education (especially in the Charter School movement) and continuous acrimony and division over public school performance (Goodenough 2001; Wilgoren, 2001; Alvarez, 2001.; Schemo, 2001, 2001a; Bello, 2000; 2000a; Greenhouse, 2001). But the accountability systems had at least brought about awareness of the topic of formal accountability and had stimulated the beginnings of involvement by some of these stakeholders (Jennings, 2001; Rothstein, 2000; Wilgoren, 2000).

There had been no systematic analysis of stakeholder satisfaction. The project explored the development of criteria for stakeholder satisfaction/dissatisfaction of an accountability system by using the Texas accountability system as a starting model. Such information about how persons felt about the accountability system (and why) could provide insights to be used to modify and improve the systems.

Rationale for methodology

An extensive review of the case study methodology by Hussey and Hussey (1997) and again by Tellis (1997) points out its long history but noted the tendency for it to wax and wane in favour. Case study could provide a rich mine of information but there was concern because the results were not generalizable. This was not a handicap here since it was intended to be

more informative about theory and descriptive of specific observations of Bermuda. The results were not highly important as particular values for comparison or contrast with other cases. This case study would reveal insights that could be applied to the theory and thereby equally to Bermuda, to the Texas case and elsewhere. In any event, Tellis (1997) cited Yin and Hamel as substantiating that even a single case study need not be problematic and could be very useful.

An opinion survey tool seemed an appropriate method of gaining information about the public's attitudes towards the education accountability mechanisms. Referring to the research questions, it was desirable to find out whether or not educators had different opinions on this matter from others. Also, what differences might distinguish these opinions? Since no direct example of such a tool/study was found in the literature it was appropriate to develop one.

The public opinion survey is a very familiar measurement technique that is designed to explore the subjective feelings of the public about an issue so as to inform and guide policy makers (Fowler, F. J. 2002). A survey is a scientifically sound method to produce quantitative and numerical descriptions about the study population. Fowler (2002) follows Dillman (2000) in referring to the 'total survey design' perspective as the intent to follow the 'best practices' in the three main focus areas of sampling design; questionnaire design and the data collection methodology. Strategies in each of these areas have been well established and refined since the 1950's and although there is debate about the details and best methods for a particular case, the relevant principles are well understood (Fowler, F. J. 2002; Alreck and Settle, 1995; Salant and Dillman, 1994). A survey can provide data that are not available otherwise. By attention to sampling and question design and collection method a survey can provide clear and meaningful data that can be extrapolated to a much larger population with confidence and with substantial savings of manpower, effort and economic costs over other methods. The opinion survey is also a favourite customer-centric tool of the quality movement.

Description of the 'case study opinion survey' site

Bermuda is an island nation of about 60,000 persons and very sophisticated education needs based on a knowledge economy (Bermuda Government, 1999). Its mixed K-12 public education environment is built on the British model but the teachers are primarily trained in North America (Bermuda Government, 1999; Bermuda Government, 1996). The upper levels of the public school system and the college level show heavy American influence through the following:-

- Most of the educators have been trained in Canada and the USA.
- Many American contract teachers serve in the local school system.
- Most Bermudians go to the USA or Canada for tertiary education.

Additionally, Bermuda and the USA share sport, cultural and family ties. Yet Bermuda is steadfastly distinct with a unique blend of American, British and Caribbean influences.

Benefits of the 'case study' location

The author was the sole project operator and therefore some possible problems of inconsistent personal bias or influence were avoided. The author had sufficient status within the Bermuda education community at large to secure the willing involvement of a comprehensive array of participants. He had worked 18 years at Bermuda College and had taken all his own pre-university education in Bermuda. He also had served the Bermuda College and the local education community in a variety of roles such that the education system and past and current players were easily accessible. The author also had experience of interviewing and surveys and focus groups as regards local education research. These were all strengths for case study design and operation (Tellis, 1997).

The Bermuda population was very accustomed to being surveyed for public feedback on a variety of topics. The year 2000 Census had just been completed and an island wide Literacy Survey Pilot Project was underway in Summer 2001. Logistically, all areas were within less than an hour's drive of the survey base. The small population and greater intimacy produced a willingness to respond and to value the chance for input.

Education was in some flux in 2001. The US education reform movement had also impacted Bermuda and by the early 1990's the Education Planning Team (EPT), a local reform vehicle, had recommended a substantial restructuring of the public K-12 system to secure a more equitable system to produce superior student outcomes and education quality. Another American-based international consulting group had reviewed the management processes and the curricula of K-12 education in 1996. (This had essentially been an inputs-based accountability assessment measuring of the system.) There was therefore much interest in Bermuda about school accountability systems. Yet, because of the differences between the Bermuda and American systems in general, participation in a case study research project did not overtly threaten any vested interests.

Questionnaire Design

Following recommended research protocol for question design (Hussey and Hussey, 1997; Tellis, 1997) interviews were first held in Texas in Spring and Summer 2001 with a selected number of persons, who could give insights from the various perspectives of student, politician, education administrators, employers, etc. about education objectives and accountability systems in general. This directed attention to what information might emerge in a deeper and more comprehensive inquiry and permitted the design of appropriate questions to elicit such responses. These interviews included open-ended as well as structured questions and were both by questionnaire and some face-to-face interviews. They were themselves designed on the basis of information and perceptions gleaned from the literature, the press and the electronic media as well as discussions with education colleagues about the Texas and other education accountability systems. This yielded material for question scripts for a structured instrument that was delivered by hand and electronically in a pilot test to groups of colleagues available through the School of Education at the University of Texas at Austin in Spring and Summer 2001.

Questionnaire development

By pilot testing American audiences confirmed that the question format was comfortable and the delivery system was feasible. They also gave useful information for refinements to the questions so as to better secure responses that would be understandable; unambiguous;

complete; able to be disaggregated or clustered; able to be quantified; and able to be made visual or graphic. Careful attention was given to the instrument design (Fowler, F.J. 2002; Alreck and Settle, 1995; Salant and Dillman, 1994).

The final questionnaire that was used in the survey is shown in Appendix A 1.1. Respondents answered questions in three sections. One section concerned demographic information. Another section probed the respondent's views on the purposes of education. This section was to "warm up" the respondent to thinking about education and associated accountability issues. These two sections together occupied one-fifth of the questionnaire. Four-fifths of the questions required the respondent to provide ratings for a number of comments about or determinants of and criteria used in outcomes based-accountability. The material for the questions either originated from items currently in the Texas K-12 school accountability system or from modification of quality assessment and accountability tools in TQM and CQI in business and institutional settings.

Of the 4 sections, Section 1 required general information on the respondent, including his/her educational background. Rose and Gallup (1999) suggested that educational background influences the attitudes that are held towards the public school system. Section 2 explored their views on the major purposes of education. Respondents rated the suggested purposes of education on a scale of 0 to 5 (0 indicates not relevant; 1 indicates a least important view; 5 indicates a highest importance view) and had the opportunity to give some open-ended responses on the purposes of education. Section 3 involved giving grades responses (scaled 0 to 5 as in part 1) about rating school attributes. Section 4 similarly provided scaled responses about attitudes on accountability and issues for school improvement.

Sampling and administration design

Although classic random sampling strategy would allow the use of the powerful tools of inferential statistics in the analysis it was considered needlessly difficult in this instance where economy was important. It required generation of random numbers; the application of such numbers to the population and identification of the respondents; and then securing the responses. This was needlessly complex. Instead, stratified random sampling was used. The

original research hypothesis was that educators and the general public would view outcomes-based accountability issues associated with K-12 education in distinguishably different ways. This determined the different strata or respondent groups. In this format, a simple random sample was taken from each stratum. This required knowledge of the size and general features of the potential respondent populations. A relatively large sample size (one that exceeded 40 respondents) permitted the use of simpler applications of inferential statistics. The process was relatively easy as target populations were small (under 500 individuals). This also kept costs and general administration needs low.

Close knowledge of the research environment produced a suitable process for a random distribution of the questionnaires. A table was set up at a conference of local educators and the survey forms were offered by hand to potential respondents at various times over a two-day period by the researcher. This conference was a major activity for local educators and the great majority of appropriate persons attended. It was not as easy to access the general public by the same method. Instead, after obtaining the support and approval of umbrella groups such as the major trade unions and informing the public-at-large by community radio programming, several distribution points were set up and manned by research assistants. These persons included shop stewards at their work sites; a church congregationist and also a local health service provider.

Therefore, participants were a random sample of educators and parents from a restricted pool (educators who served in the public system and parents who worked in the public or in the tourism/hospitality sector). They responded via a self-administered questionnaire with scaled answers. About 300 forms were distributed with the intent of securing about 200 useable responses. This was within the returns rate that was forecast by Dillman (1978). The numbers reflected empirical considerations of the importance of the accountability system evaluation to the various stakeholders as judged by the author. All forms were numbered and all questions were coded to assist in later analysis.

Reliability and validity

The reliability and validity of the research study are important for the reader and the researcher. Validity reflects the value of the results. It does not belong to the samples, measures or the design of the research (Trochim, 2001). However, the processes of the research must impact the information from which valid inferences/conclusions arise.

In positivist understanding we have an absolute value for data and reliability refers to how often the researcher will observe that 'true' value. There, validity is measured by how often that observed value is substantiated by other independent measuring systems. In non-positivist research there is no ultimate truth. Data is not unchanging and can lead to varying inferences/conclusions. Validity is "the best available approximation to the truth of a given proposition, inference or conclusion" (Trochim, 2001). Reliability, according to Trochim (2001) "is the consistency of [the] measurement or the degree to which an instrument measures the same way each time it is used under the same conditions with the same subjects". Reliability is about stability and consistency. It does not mean the values are correct and will yield valid conclusions/inferences. Valid conclusions must come from 'correct information,' however.

The reliability of a study is determined by the extent to which (1) its component features in the total design are consistent with one another; and (2) it can be replicated and yield similar outcomes/inferences/conclusions. Good external reliability for a survey is due to proper interaction between questionnaire design, sampling, data collection and analysis. It is shown by the extent to which the total design concept is applied, documented and can be replicated. Internal reliability is related to internal consistency and applies when several observations are made to obtain a score for each subject (Graziano and Raulin, 1993). Generally, the more observations that are made to obtain a score, then the more reliable the score. The questionnaire design used here allowed several items, which were thought related to the respondent's opinion, to be asked and furthermore each item was itself addressed by several variations of the question. Babbie (1999) states that self-administered survey research is strong on reliability. Reliability can be computed statistically as an indicator such as Cronbach's alpha. This compares the consistency of response for all items on the survey and

may be computed for each sample. This would need data from previous similar studies and hence was not useable here as this was a first study of its type. Attention to the total design concept assured good reliability.

High reliability is necessary for high validity but does not by itself establish validity. Validity for an opinion survey is highly debatable by some (Babbie, 1999) if it assumes some absolute truth or true concept. But an opinion is subjective. Validity is not determinable by mathematical methods but it is estimated. The researcher can take steps to make the results more likely to be valuable. Trochim (2001) suggested that validity was enhanced by (1) good reliability of the measuring tool (2) good implementation protocols (3) good statistical power. Good reliability of self-administered opinion surveys has been mentioned earlier. The clear and consistent methods applied gave good implementation. An attempt at good statistical power was made by having a large sample size (about 160 respondents) with near equal numbers in the groups to be compared. Steps can be taken in the planning stage to provide for enhanced validity but since validity is a feature of the conclusions/ and inferences it can really only be estimated at the end of the study.

Securing overall study quality

This plan of research was better than that laid out as an exemplary core design for case study by Tellis (1997) and exceeded requirements proposed by a branch of the Organization for Economic Co-operation and Development (OECD) for a multi-national educational case study in 2000. The research began with a pilot study in Austin using a questionnaire that was developed from a theory base and after discussion with persons who were knowledgeable about relevant aspects of the Texas accountability model, such as Dr. Wilhelmina Delco and several principals and superintendents in the Texas system. In the pilot this first draft questionnaire was tested with a small sample of university students and education professionals (mostly administrators) to evaluate particulars such as clarity, length of questions, relevance of items, vocabulary, etc. Then came adjustment to the questionnaire to a form that was the final draft to be used in the field research.

The final study population was cooperative and the context for the survey was made even more hospitable by specific actions of the author like a pre-launch radio discussion to build enthusiasm and doing the initial distribution via a local educational conference. Various particular attributes of the case study situation such as ready access to stakeholders, lack of conflicting interests about accountability models on behalf of respondents, and the convenient logistics were also important.

Data Analysis

Data were obtained as respondent's answers to three types of questions. The first set of questions (Type 1) concerned demographic information. These are questions 1 to 5. The answers to Questions 6 to 15 (Type 2 questions) were the respondent's opinions about the values and objectives of education systems. Questions 16 to 71 (Type 3 questions) were of features often used in ratings for outcomes-based accountability systems in education.

Data for Type 2 and Type 3 questions consisted of estimation of the importance of each factor in turn on a response scale of 0 to 5. A value of 0 indicated the factor was considered irrelevant and 5 indicated it was very highly important. Type 2 and Type 3 questions were survey content questions. (The non-numeric descriptive data available from Q.4 was used as an internal check/verifier of other demographic data. It spotlighted any incongruent data about educational background in the data cleaning process for questionnaires. Any with seriously incongruent/flawed information were discarded.)

Computed mean ratings were calculated and the numbers of respondents also noted for each survey content question. (For use of the SPSS system the responses to the numbered questions were converted into an alphabetically derived identification format. Answers to Questions 1 to 5 were noted as A to G. Answers to Type 2 questions were identified as cases H to R. Data for responses of question 16 to 71 were noted as S to BV). Mean ratings for different questions were compared as were means for the same question as answered by the Educator and Public group.

The pattern of the data as shown by the responses for a population group is as important as the numerical values that are used for the descriptive statistics. Note that in Table 3.1 all groups have a mean of 2 and clearly different data patterns. These include a homogenous set of responses of 2 each; a set of two very disparate pairs (one pair gives a high rating of 4 each while the other pair shows scores of zero) and other possible results.

Table 3.1. Table of imaginary response scores of four respondents in a group where each set of results has the same statistical mean of 2. The response patterns as shown by the data are sharply different in some cases.

Score of first respondent	Score of second respondent	Score of third respondent	Score of fourth respondent	Mean of respondents' scores
2	2	2	2	2
3	2	3	2	2
4	4	0	0	2
5	1	1	1	2
3	3	1	1	2

Differences in data patterns were shown by plotting frequency histograms of the response scores for each question. A frequency histogram of the response data for each question for the educator population was matched against a frequency histogram of response data for the same question for the general public. One was overlain on the other and compared visually for similarity. They were judged as similar if they had the same general shape or represented the same or similar portions of a normal frequency curve.

Response pattern

Each response was a unit of data. Slightly more than 80 respondents made up each population sample. The 71 numbered questions could yield 75 data pieces per respondent. The response option ‘no answer’ indicated intents ranging from a lack of understanding of the question; deciding that none of the available options was appropriate or simply forgetting to supply an answer. Note that zero (0) was used if the question was considered irrelevant.

Survey content questions were checked for the pattern of whether answered at all and what type of answer.

Limits applied to respondents

If stakeholders are anyone on whom the education system impacts then the net extends far beyond just students and teachers. Parents must be included. The stakeholders considered here are limited to those indicated in Table 1 of the Appendix. The main focus is on responses by parents and educators. These would in the author's experience and opinion have the most to offer in an investigation into K-12 education accountability.

Limits on the number of responses handled

Limits on the number of responses that would be analysed were due to the scope of the project as initial research. This total and its distribution were to provide enough respondent answers in each class to allow for a broad range of opinions while minimizing the biases that easily occur when using only few respondents. The total number of 200 to be analysed could be handled in the time and resource constraints of the project. Larger numbers of responses to be handled would be very onerous. A suitable response was determined as one, which had answers to at least half of the questions in each of the sections 2, 3, and 4 of the questionnaire and which (preferably) was complete in answers for questions of section 1. A limit on the time to collect responses was set initially at 4 weeks but the majority of responses were attempted to be collected within one week.

Summary

Because the case study does not attempt to mimic the Texas scenario, it could have been carried out in a number of different locations and still have offered the chance to contribute to the theory of how various stakeholders could be enabled to assess the accountability systems of K-12 public education in various states in America. Texas was used as a base model system as it had a well-known reputation in America and it was well documented. Using Bermuda as the scene for the case study allowed the use of certain resources that the author and the situation held that could be strategic to its success. Bermuda was in a state where the project could be expected to attract willing participation of the required

stakeholder groups. The author had access to the necessary breadth and number of contacts to make the exercise possible with the time and other resources that were available. Public opinion surveys are a familiar investigative tool and the methodology is well understood and documented. It was a suitable research methodology within the parameters of the questions being asked and the logistics of the situation.

Chapter 4 – Results

Introduction

The hypothesis is that "educators will have a significantly different opinion (rating responses) of the estimation of the importance of factors concerned with outcomes-based accountability to the opinions (rating responses) of non-educators in the general public". The results were the answer to the original research questions. The results should reveal if it was possible to use a customer-centric survey type questionnaire to allow persons to express their opinion on the education accountability system. The results should describe the opinions of the various groups and indicate their degree of similarity and distinctiveness. They would also indicate the more influential opinion descriptors.

A self-administered survey questionnaire was randomly given to members of the general public and classroom educators. The original response period of four weeks was eventually extended to 12 weeks to secure enough respondents. A few persons (less than six individuals of the perhaps 300 approached) refused involvement but many more took the questionnaire and simply never returned it. It proved difficult to survey sites concurrently and intervening holidays were quite disruptive to getting returns. Respondents were anonymous but returned questionnaires were coded for later identification if needed. Virtually no unusable responses were returned. A few cases had incomplete results, which appeared to be due to the respondents simply forgetting to complete a page. Persons often omitted responses to two particular questions; i.e., Q 14 and Q.15.

Return rate for survey forms

More than half of the educator responses were from conference attendees. Persons who did not return the forms by the conference end were either difficult to track or never bothered to turn in the form at all. The overall return rate was 33% (82/250). The return rate for forms given at the conference was 38% (53/160) with most returns being immediate. The return rate of the forms from school sites was 13% and this was after extending the original participation period. College educators responded more frequently than the K-12 teaching fraternity (40% or12/30 as compared to 17% or 17/100) when approached via a general delivery format but took as long to respond.

The extended response period for educators and the extra administration caused a delayed approach to the general public. A random survey of the public was used within a stratified or selected sample with the assistance of the two major workers' organizations in Bermuda --- the Bermuda Industrial Union (BIU), and the Bermuda Public Services Association (BPSA). This gave the full range of employment categories, age and educational background. An associate assisted the process at four hotel sites and three other workplaces. The overall response rate from the public was 36% (90/250).

Table 4.1 Survey administration and response rates for various stakeholder groups.

	Maximum population	Targeted sample (portion of whole)	Responses	Actual rate of response	
Educators	About 600	250 40%	82	33%	
General public	About 60,000	250 0.4%	90	36%	

Response pattern

The response pattern was noted for the various groups of questions according to question type. Scored answers were obviously important as was whether or not a question was answered as well as distinguishing between a non-response and a zero answer.

In the group Q.1 – 5 there were about equal numbers of non-responses for both populations. In both most non-responses occurred in Q. 5. Since Q.5 yielded usually at least 2 data units this represents about 4 persons in populations of more than 80 respondents in each case. As regards Q.6 – 13, fewer non-responses were seen even though there was a greater number of potential data units. This non-response rate represented fewer respondents than for the first group. Educators answered every question in this group and never considered any of them as non-relevant whilst the general public gave 3 instances in 640 of a zero answer.

Table 4.2 Pattern of Responses for Groups of Questions for the Different Populations answering the Questionnaire Survey. (The number of instances 'N' in each question grouping is shown for the total sample population.)

		Q.1 - 5	Q.6 - 13.	Q.15	Q.16 - 44	Q.45 - 71	Q.16- 71
No answer	Public	17	14	74	67	60	127
No answer	Educator	14	0	86	41	30	71
Zero as answer	Public	Not apply	3	Not apply	102	73	175
Zero as answer	Educator	Not apply	0	Not apply	25	51	76

(The response rate to Q.14 was lower than 20%. These responses were free responses and were not further analysed in this research.)

For Q.15 there were about 240 data units per population. The public gave 74 no responses and the educator's tally was 86. Yet educators responded very diligently to the immediately previous question group. Such high non-response in each case (about 30%) was clearly a deliberate non-response pattern for each population for Q.15.

When considering Q.16–44 (which represents about 2400 data units as there are more than 80 respondents and 29 available data opportunities for each) there were 67 non-responses for the public sample and 41 non-responses for the educator sample. These were less than 3% of the missing responses in each case. The group Q.45–71 had slightly fewer potential responses but had even larger decreases in the numbers of non-responses. Here non-responses had even less impact. The non-responses for the public represent 127 instances of a potential of about 4500 data units and 71 non-responses for the same potential in the educator sample. Non-responses were about 2.6% for the general public and about 1.6% for educators for the data from Q.16 to 71.

For the group Q.16–44 the public recorded 102 zero responses while the educator sample had 25. The public noted items as non-relevant four times more frequently than did the educators but even so the 102 zero responses was only slightly more than 4% of the possible responses. Similarly for the group Q.45–71 the public had 73 zero responses and the educators had 51. These were about 3% and 2%, respectively, of the total responses. It takes a deliberate act to record a zero response but a non-response may be unintentionally due to forgetting to answer or overlooking the question accidentally. On average every respondent recorded 3 zero responses in the public sample and between 1 and 2 zero responses in the educator sample. Educators were much more likely to find questions they considered not relevant in the Q.4–71 group, but for the public it was the converse and they more frequently recorded zero responses in the Q.16–44 group.

In summary, very few questions were ignored and the vast majority was answered. Considering questions 16 to 71, there were 198 cases of No Answer of about 9,500 possible answers. This was just over 2 % of possible responses. Cases where zero was given as an answer numbered 251 in this same question range. This is between 2% and 3% of the total possible answer pool. Overall, the response patterns show that the questionnaire was very effective at engendering an answer (about 98% of cases) and generally the answer was not one that noted the item as being non-relevant.

Data pattern

The frequency histograms of the matched question responses showed much general similarity when matched visually or overlain. Individual questions gave very different frequency distributions. Twelve (12) of the 56 frequency histograms of matched responses for the two populations were noticeably different (See Appendix D.) Table 4.3 has relevant descriptions/explanations of these histogram differences. Most differences reflect response data of the two populations representing different parts of the normal curve. For six of the twelve pairs of frequency histograms the curves were displaced to the left or right with reference to each other. The remaining six instances fitted no simple explanation such as a displacement of the curve to the left or right.

Table 4.3. Comparison of descriptions of the frequency histograms for the instances where sharp differences were seen by visual comparison.

Identity of matched question response	Educator population results	General public population results	Simple explanation of possible relationship
S Q.16	About 3/4 of the left side of a normal curve	One graph is central area of normal curve at low values; other is central part of normal curve at high values	No simple explanation
U Q.18	Left I/2 of normal curve	Flat graph at low values; flat graph at high values	No simple explanation.
V Q.19	Normal curve with high level of low values and of Zero responses	Right half of normal curve with very large level of Zero responses.	General public curve represents a leftwards shift
AD Q.27	Left 2/3 of a normal curve	Left 1/2 of normal curve	Educators curve represents a leftwards shift
AK Q.34	Left 1/2 of normal curve	Flat curve for low values; inverted central portion of normal curve for high values	No simple explanation
AO Q.38	Left 1/2 of normal curve. Several Zero responses.	Left 2/3 of normal curve. Sizeable Zero category	General public curve shows leftwards shift.
AQ Q.40	Consistent low responses including Zero responses; near even set of high responses	Left 1/2 of a normal curve with high frequencies of two low categories and of Zero responses	No simple explanation
AR Q.41	Left 1/2 of normal curve with two peaks at lower end	Flat response for lower values; right hand 1/2 of normal curve at high values	No simple explanation
BF Q.55	Left 2/3 of a normal curve	Left 1/2 of normal curve. Some distortion at low end.	General public curve shows a leftwards shift..
BL Q.61	Left 2/3 of a normal curve	Left 3/4 of normal curve	Educator curve has similar pattern but looser spread of results
BS Q.68	Left 1/2 of normal curve. Distorted low end	Flat curve but a central peak	No simple explanation
BT Q.69	Left 1/2 of normal curve	Left 2/3 of normal curve	General public curve shows leftwards shift.

Describing the Data

Data for Type 1 questions were nominal numbers that identified categories. For Q.1 numbers represent age bands. For Q.2 numbers identify the gender. For Q.3 numbers indicate the site of pre-college schooling. For Q.4 the occupational category could be transformed into a numerical identifier of the standard classification of occupations. For Q.5 up to 3 numbers indicated the academic background by supplying information about the highest level of formal education and any training and its timeframe.

Type 2 and Type 3 question responses were mainly on an interval scale of 0 to 5. (The response to Q.14 was a word description. The number answers for Q.15 were nominal. Data from Q.14 and Q.15 were not handled further.) The original data from respondent returns were collated as appendix C 1.0 and 1.1. For all questions of Q.6 to Q.13 and Q.16 to Q.71, the answers were interval numbers with statistical descriptions of the mean, median, mode, standard deviation and variance. The size of the samples allowed the use of statistical formulations that approximated the results being normally distributed.

Differences exist between the means calculated for responses to various questions as answered by the same population. Sometimes these differences in ratings for the various responses appear slight and at other times are fairly large (about 0.5 or greater). When the responses given by the two populations for the same question (type 3 questions) are compared in pairs, differences are again seen. The descriptive statistics for the data are given in Tables S1 to S28 in Appendix C1.0 and 1.1. The comparative means are shown as bar charts in Appendix B.

In the complete display of the descriptive statistics for the two sets of population response there are instances where the matched responses show response means that are identical or have extremely small differences. Yet other instances show some differences in values. After inspection, the researcher decided to calculate the Z statistic for those matched pairs where the difference between the means for the responses of the two populations was at a value of 0.40 or greater. The computed values are given below in Table 4.4.

Table 4.4 Determination of the Z statistic and the related significance for the difference in means of matched pairs of responses from the two population samples for those questions with a difference in means of 0.40 or greater.

Identifier for paired means	Difference in paired means	Z statistic	Determination of significance
T Q.17	0.40	2.20	Significant at 99% level of certainty
U Q.18	0.74	3.49	Significant at 99% level of certainty
V Q.19	0.69	2.91	Significant at 99% level of certainty
AA Q.24	0.62	3.48	Significant at 99% level of certainty
AK Q.34	0.92	4.04	Significant at 99% level of certainty
AO Q.38	0.48	2.29	Significant at 95% level of certainty
AP Q.39	0.93	2.21	Significant at 95% level of certainty
AR Q.41	0.82	3.48	Significant at 99% level of certainty
AV Q.45	0.59	3.34	Significant at 99% level of certainty
AZ Q.49	0.45	2.23	Significant at 95% level of certainty
BB Q.51	0.40	2.48	Significant at 99%level of certainty
BH Q.57	0.41	1.81	Significant at 95%level of certainty
BL Q.61	0.39	1.68	Significant at 95% level of certainty
BR Q.67	0.64	2.97	Significant at 99% level of certainty
BS Q.68	0.62	2.58	Significant at 99% level of certainty
BT Q.69	0.86	4.43	Significant at 99% level of certainty

The results show that in 16 of the 56 pairs of matched means, the responses of the Educator and Public groups show significant difference at a confidence level of 95% or better. There

is a 5% chance or less that the results could be that different by chance. In fact in 11 of the cases there is only a 1% chance or less that the differences shown in the results are not significant.

Comparison of means of indexed scores

When the means of the original numerical scores are compared directly the range is very narrow and therefore discrimination is not easy. This readily shows in the visual comparison of the means by the bar charts (See Appendix B.) as referred to earlier. A clearer picture can be given by developing a system with a wider range for the results. This was done by producing a weighted score or indexed score. The equivalent of the weighted score for the original numeric score is shown in Table 4.5. The weightings are subjective but a range of zero to one hundred is used as this presents a form that is very familiar in everyday routine in the comparisons of percentages.

Table 4. 5. Original respondent score and weighted equivalent to form the indexed score.

Respondent score	5	4	3	2	1
Weighted score	100	75	50	25	0

The full set of indexed scores that result is shown in Appendix F 1.0. The mean scores are produced for each item that is noted as the responses of Q.16 to Q.7 and a rank order of the score is also noted. Where the mean indexed scores differ by 7.6 or more between the Educators and the Public the result is significant. Comparison of the mean indexed scores indicates 17 instances out of 56 possible pairs where there are significant differences between the two populations. The compared mean indexed scores and the relevant question identifiers are shown in Table 4.6.

Table 4.6 Comparison of mean indexed scores for Educator and Public samples where a difference of 7.6 or greater is significant. The relevant question item is identified by number and alphabetic code in each case.

Question	Educator score	Public score	Difference		Question	Educator score	Public Score	Difference
18 U	75.3	64.1	11.2		50 BA	44.9	55.0	10.1
19 V	41.1	31.9	9.2		51 BB	77.6	86.3	8.7
21 X	85.8	77.7	8.1		57 BH	68.2	59.6	8.6
24 AA	82.9	71.0	11.9		61 BL	71.7	61.5	10.2
34 AK	79.0	59.4	19.6		66 BQ	74.0	82.0	8.0
35 AL	80.2	71.3	8.9		67 BR	29.2	44.3	15.1
39 AP	79.1	68.8	10.3		68 BS	62.3	53.4	8.9
41 AR	62.0	51.4	10.6		69 BT	82.6	63.0	19.6
45 AV	79.3	89.9	10.6					

Factor analysis of data using SPSS

Factor analysis is a complex but useful statistical tool that was employed both to reduce the data and clarify the structure of the large data mass. The data for Q. 16 to Q.71 consisted of up to 56 responses by each of about 160 respondents. All this data resulted from those questions on the survey item that were intended to specifically inform about the respondent's opinions on K-12 education accountability.

Factor analysis showed that the questionnaire survey items could be grouped into 17 categories statistically as defining the opinions. Categories or factors are determined by the amount of variance in the data that is explained such that a few factors usually explain the bulk of the data and successive factors are responsible for progressively less of the variance seen. The number of factors is determined subjectively by reference to the theory underlying the study. The first 10 factors were used as the base since together they could explain approximately 56% of the variance observed.

Using the 10 factors as a substitute to describe the data from the respondents it became easier to look for relationships between the respondent's demographic features and the data that described their opinion by the responses to Q.16 to Q.71. The former was cast as the Independent Variables (IV) and the latter as Dependent Variables (DV). The results were analysed using Oneway ANOVA. Full results are given in Appendix H. Instead of the actual data as DV the data were collapsed into 10 factor scores. Each factor score is determined by the data from the questions that contribute to or compose that factor. These are seen in Appendix H.

In some cases, ANOVA tests showed that there was a significant relationship between the factor scores and certain of the Independent Variables. These are shown in Table 4.

Table 4. 7 Summarized results of ANOVA tests showing significant relationship between the factor scores (DV) and the demographic data (IV). Extracted from the full results shown in Appendix H.

Independent Variable	Dependent Variable	Significance Level
Respondent Type	Factor 4	0.025
Respondent type	Factor 5	0.000
Gender	Factor 5	0.008
Last educational event	Factor 6	0.026
Site of K-12 education	Factor 10	0.027

All levels of significance below 0.05 denote a meaningful relationship or interaction.

Chapter 5 – Discussion of analysis and interpretation of results

Introduction

The broad successes of the research project are possibly best demonstrated by an in-depth analysis and discussion of the interpretation of the results and the conclusions/inferences that flow from them. This will begin with a review of our collected data or descriptive statistics in the context of the underlying theory. Further review of the data reveals patterns that support further inferences that will be posited within the framework of current theory and the research. Finally, the study will be critically examined as a whole to suggest its significance and contribution to the literature and what might be subsequent steps or actions to build on this.

Reliability and the return rate

Reliability was addressed in the Methodology section but must be also considered after the fieldwork. The design of the study and its execution were planned so as to maximize reliability; i.e., to maximize the chance that if the study were repeated as described and under the same conditions, then the same general data would be obtained.

Opinions are not static and are expected to change in the long term (as the respondent's life experiences and casting framework for the opinion change with time) and are quite fluid in the short term (because an opinion is constantly bombarded with change agents). Survey research is generally high on reliability and increases when self-administered (Babbie, 1999) since social constraints and influences are removed because of anonymity and the feeling of self-control of the respondent. If care is taken to address the necessary elements of tool design and application according to the 'total design concept' then the respondents will be able to accurately and easily express their opinions. Language that is difficult to understand; tedious or clumsy formatting that is not easily used; complex items that are vexatious and other design and administration flaws may show in both a low return rate and a poor response pattern of those received. The former means that persons did not often bother to respond. The latter might show as many incomplete questionnaires and responses with much modification and comment by the respondents to note their dissatisfaction with the study.

The return rate of questionnaires in this study was 33%. It is very difficult to predict an appropriate expectation (Fowler, F.J. 2002; Babbie, 1999;Heiman, 1998) since this varies with so many features such as the topic and the interest level of respondents; incentives and publicity; competition from other activities as related to weather, season and other factors, and numerous other variables. For a self-administered survey it can be higher than 70% or as low as under 25%. A specific group may be targeted and incentives, prizes and other inducements used to increase participation. Very broad and high impact media coverage and communication may be added. Despite the very simple nature of this study in terms of manpower and funds and inducements a creditable return rate was achieved.

This is probably related in part to features of the Bermuda situation that were cited as benefits of the case study location. Bermuda is a small but sophisticated community where there is keen interest in education and a strong appreciation of 'world class business behaviours'. The Bermuda economy is sometimes referred to as Bermuda, Inc. to reflect the belief that the business lens has dominated the past 20 to 30 years of Bermuda's history. Although it has some decreased prominence now, it is still a powerful perspective. The high living standards seen in Bermuda are built on capital intensive and low labour demand but high-end business. These must be underpinned by an efficient and effective education system to provide the pool of high calibre/highly educated persons who must master the economic engine.

It is also interesting to note that the educators and the general public had comparable rates of return. It was easier to enlist the cooperation of the educators (probably because the educators would find the opinion topic more immediately relevant) but once engaged the general public seemed to have at least a similar level of interest as indicated by the return rate. Educators would be expected to have encountered various aspects of the accountability mandate and the consequent accountability systems at overseas conferences; in reading the literature and professional articles; and in discussions with overseas colleagues, etc. There has not been the same attention formally to this topic in Bermuda. In Bermuda, as elsewhere, the general public has not been accustomed to being able to present opinions or feedback on education and have this formally considered by the 'education experts'. Education is one of

the few services areas where the customer's view and right to comment has been routinely dismissed. Indeed few educators will accept that the customer concept is acceptable. It therefore will take more than a single study to alter this status quo. The Bermuda public is often polled on all kinds of matters. There are routine telephone surveys and polls and surveys with varying levels of underlying professionalism. There are several reputable local organizations and government units that gather opinions to provide background and support for policy analysis and decisions in the government and commercial sectors.

Reliability and non-responses

Analysis of the type of respondent and the respondent pool is also important to reliability. The smaller the pool of respondents, then the more likely it is to be biased. A small pool may represent persons who are drawn to respond because of some atypical interest in the topic or may indicate that the study process had some limitation that prevented the broadest base of respondents from participating. The average educator respondent was female, 35-45 years old, had a Masters degree and fitted the author's expectation for this case study. The profile of the general public respondent/parent respondent was also agreeable with information from the Report of the 2000 Census of Population and Housing (Bermuda Government) and the author's general expectations.

Educators may be more interested in educational accountability than the general public. They may be more knowledgeable about the details and issues as it may possibly more closely relate to their daily life and work. They may feel it is more personally important for them to respond and attempt to influence and impact the collected data on stakeholder opinions. Educators may be more likely to respond to survey questionnaires; may find them easier to handle because of the educator's generally higher educational level and may tend to be more assertive in public. As a group the general public may be viewed as displaying the converse features. Also some respondents among the general public may not have children in the public K-12 system or may not have used it themselves and may be less able to give informed opinions. Also, there was a predominance of persons represented by the BPSA in the general public category of respondents. These would be the more educated persons as the

BPSA especially represents the civil service workers and others who would be expected to be in the higher academic levels.

There might be a shortfall of less educated respondents. The Bermuda population is deeply questioning the comparative merits and structure of the public and private education systems. There has been a tacit agreement that education issues were not fair game for party politics and there has been often frank and open contributions from all sides and many voices. The researcher is respected and well regarded such that the study has been welcomed and supported by all. The system was not designed to detect any bias in the respondent pool but the simplicity of the questionnaire design and its convenience to complete (as reported by numerous respondents) together made it amenable for all.

In Bermuda, for historical and socio-economic consequences that emanate from colonialism and racial bias that were common throughout this region, persons of colour use the public school system more often than they will use the options of private and charter education. A question to be asked is therefore whether this continuing preference is linked to different perspectives about the efficiency and effectiveness of K-12 public education. Do they have different opinions about the accountability and opportunities for improvement and success of the systems? Recent work suggests that socio-economic status is a potent factor in how persons deal with accountability issues in public education. (Skrla and Scheurich, 2001). Perhaps a race identifier should be included and greater attempts to target less educated persons as respondents.

Reliability and the response pattern

There was a very low incidence of persons who did not answer all the questions and there was only a low incidence of questions being noted as irrelevant or not important. Even if persons are somehow persuaded to participate, they must be able to use the tool appropriately to enable the researcher to get meaningful data. Instructions must be followed correctly so as not to produce dubious or muddled data. Respondents must be sufficiently comfortable with the language and format to feel that it allows them to express their opinions. They must feel the items used are appropriate and adequate or else they may tend to add their own or alter

the instrument if they bother to respond at all. There was only one case of a respondent who added comments to the instrument and this was to express a stronger view than that indicated in the tool.

The survey instrument and procedure were pre-tested by first interviewing persons in Texas; extracting indicators based on those from the Texas accountability system and others in the USA; and then piloting a version on a population from the School of Education of the University of Texas at Austin. This broad design was refined by particular reference to language and practice in Bermuda as based on the author's wide local knowledge and experience.

The success of the study reinforces the advice put forward in the 'total design concept'. Good question design involves principles that can be generally applied. The Bermuda population is like many American small cities in terms of size and socio-economic descriptors. The basic instrument would therefore not be expected to be radically different but would be refined for various jurisdictions as is often seen when using a tool in the Deep South as against the North or on the East Coast versus the West Coast locales. The success as seen in the return rate; the response pattern and the lack of overt distorting bias in the respondents point to a high reliability such as was expected according to Babbie (1999).

Validity of the study

Validity is a feature of the conclusions/inferences of the study and is "generally" weak for self-administered opinion survey research according to Babbie (1999). He notes that people's opinions seldom take the very structured form that is required by the response constraints of the survey questionnaire. They are constrained by the vocabulary and form of the items; they are similarly constrained by the allowable responses; and by the overall form of the tool. Yet one's opinion exists of itself and is independent of any survey tool. An opinion is not thought of in relation to a specific indicator but that may be what is used to express that opinion according to the tool. The response that is given must therefore be interpreted as an approximation of what the respondent intended or would have really said if given absolute freedom of expression. It is also an approximation of the interpretation of

what the researcher intended to ask in seeking the opinion. This opens the whole Pandora's box of 'meaning and constructivism'. Babbie speaks of researchers who say there can be no assessment of validity in opinion research since it cannot be validated against some absolute truth because it is opinion we are seeking. It is valid if it has been produced by a methodology that has good reliability. If the KMO value (0.657) is considered then the reliability is acceptable but not exemplary. But it is worth repeating that this was a first study and surely with more resources to administer the study and a larger sample and other measures to increase the return rate, then an improvement should follow in the computed reliability. It is still clear that the success reported earlier suggests that the study is reliable.

Validity is also considered by some authors in the context of "how well a measure (reflects) /measures a concept". This can be reconciled with the view of validity being a feature of the conclusions/inferences by saying that valid conclusions must be based on valid measures. An opinion is valid because it exists of itself. Validity and the chance to be a meaningful opinion (one that is valued and respected by others and which will be used by others) must also be based on how well the tool 'communicates' the opinion. Validity is applied to the measuring tool as defined by face validity; content validity; construct validity; concurrent validity; predictive validity and criterion-related validity.

Face validity refers to peer agreement or acceptance of definitions and procedures. This was addressed in the planning and implementation stages by having the pilot and by using a methodology that is well established in other research areas.

Content validity of the tool is reflected by the extent to which the tool is validated or justified by other evidence. The use of the total design concept denotes those same features that engender high content validity. If it had low content validity the tool would have been difficult to understand and use. These would have yielded poor return rates and response patterns suggesting a cumbersome tool with many incomplete responses and noting of irrelevancy. The success in both the response pattern and the return rate support acceptable content validity. Additionally, content validity is usually scored by experts in the field. The study of education accountability systems does not have a lot of supporting literature and

indeed the tool used here is a new creation. There is a dearth of recent/current research on the Bermuda education system and its accountability issues. Hence there is no ready panel of experts and study models. But the researcher has considerable experience and knowledge of the Bermuda system. Therefore the confidence that he can express about the content validity as being acceptable should be a strong consideration. Of course the study needs to be scrutinized and evaluated by others with interests and expertise in accountability systems and especially as pertains to the Bermuda public–12 system. This will build the literature and expertise for subsequent analysis of content validity in such continuing studies.

Construct validity and concurrent validity are both tied up with the idea of being able to have a clear and absolute description of the opinion and then assessing the instrument's ability to measure all interpretations of the opinion or to show that the measures correlate with other phenomena that are associated with the opinion. Since this study is a first step and no other data exists for a defined opinion via such a tool or for related phenomena, then it means that these do not apply at his point. For similar reasons predictive validity is not relevant as it could only be so when other data have been produced in like studies.

In summary, therefore, validity of the inferences is not easily determined by external means in this study. The data are valid because the tool has high reliability and because the researcher, who is experienced and competent and knowledgeable in the context of the study and its context; has said it has high face validity; and the success in securing full and active participation suggests that it has high content validity. As the tool and the study overall have high reliability and have validity, then the data should yield valid inferences/conclusions. This links to the initial expectation as proposed in the Methodology that the data would be valid.

Inferences

Using the data the researcher is drawn to make certain inferences. These are:-

That there are significant differences detected between educators and the general public/parents as regards details of their opinions on accountability issues in K-12 public education;

That when indexed scoring is used to compare the details of the responses that describe the opinions of the two groups the scores indicate that there are some items that show the same or similar ranking in terms of importance within the opinion but there are others that are quite different;
That the opinions can also be presented as quantitative descriptive statistics including the means and standard deviations and such and an item analysis of the matched means suggests significant differences between the two populations for 16 of the 56 items.

The literature as regards the American situation suggests that there might be differences in the opinions of educators and the general public as regards accountability in education systems. Some educator groups have shown strong resistance to the concept and practice of formal accountability schemes while the public has had to embrace it throughout its general life experiences in the non-education context. They reasonably might ask why it should not apply to education. When given a means to express an opinion, then it is foreseeable that educators might be hesitant to participate or would do so and express opinions that would be different from that of the general public. Educators would probably be more negative than the general public.

On examining the items ranked as the 10 most important by using indexed scores for both populations the study results show a narrow range; meaning that there was very little difference in importance between the extremes and that they were all given high importance. The two populations shared 7 of these items and as the range is narrow there is probably no significance to any difference in the ranking order. Each group has 3 unique items. however. The educators believed it important that 'Parental attitude to school determines student motivation' and that Society's attitude towards schools needs lots of improvement'. The general public felt it important that 'Principals should be more accountable for schools', that 'Education should be improvable like other professions' and that ' Unsuccessful teachers should be retrained or removed'.

Hence, although there is agreement over the majority of which items are highly important, the items that are different may hint at and arise from fundamentally different perspectives

about education accountability. In the literature, educators do not see themselves with a role of direct responsibility in the accountability stakes for education. They did see that parents and society and many other aspects were involved and important in accountability schemes for improvement (Lessinger, 1970; Sciara and Jantz, 1972). Contrarily, the general public sees educators as intrinsically and intimately involved. The public references this to improvability that is current elsewhere in its life experiences and would hold educators directly responsible for their roles in the accountability of the education system.

As an opinion is such a complex phenomenon the survey instrument was designed to have several items that would each address various and different aspects of that opinion. Attention was paid to having items that addressed:- who should be accountable ---students, parents, educators or society.; what should the consequences of acceptable or non-acceptable performance be ---i.e., would accountability be accompanied by reward or admonition. Which would prove more acceptable --- those indicators that had been taken directly from known accountability schemes used in education or those that were adapted from customer-centred accountability schemes from the TQM world.

Factor analysis is a means of detecting structure in a complex set of data that describes a phenomenon such as an opinion. If the opinions of two groups are analysed by comparing the means of individual descriptors and there are sufficient descriptors, then just by chance we should expect some level of significant difference to be found in some of the paired means. In comparing data for 56 paired means as results from the study some significant differences are expected. The fact that as many as 16 of the 56 pairs show significant difference is itself notable, however. So many variables would be cumbersome to manipulate. It is useful as a research axiom to seek simplicity wherever possible as an explanation and only resort to complexity if it is truly necessary or offers otherwise inaccessible advantage.

Factor analysis of the data and ANOVA studies

The data were therefore examined by factor analysis. Essentially, this groups the data into patterns using correlation studies that are based on the fact that question items that contribute

in a similar way to the same aspect of an opinion should be answered similarly by a respondent. The data should show high correlation. This allowed the 56 items to be reduced to 17 and then eventually to 10 factors. These factors can be seen as components that make up the description of the opinion as the opinion is detected and defined by the instrument. Naming factors is not an easy task as it involves first identifying the question items that seem to contribute most to that factor (they have the highest factor loading) and then finding a name which unifies and describes those questions. An attempt is shown at Appendix G. This is said to be an attempt as it is quite subjective as to what values of loadings are used to decide the question items which are to be included and authors vary widely in their recommendations (Stasoft, 2003; Babbie, 1999;Pedhazur and Schmelkin, 1991). At any rate the factors are as named and shown in Appendix H.

The individual observations are grouped to form a factor score. Use of factor scores allowed further analysis of the data by carrying out an analysis of variance or ANOVA. ANOVA is often used to determine relationships between Independent Variables and Dependent Variables when population sub-groups are present. ANOVA was particularly apt since the samples were nearly equal and large (Heiman, 1998). When ANOVA analysis was carried out for the various named factors (as Dependent Variables) against the demographic data (as Independent Variables) there were some significant differences as noted in Table 4.7. These are elaborated here.

Factor 4 – Societal attitudes versus Respondent type

The original thesis behind this study was that educators and the public in the USA have different attitudes on education accountability and this would probably be seen also in Bermuda. Though less prominent here, calls for accountability are becoming louder and more frequent. The educator fraternity has been largely silent but like elsewhere they often draw attention to what others could do to improve the system and its performance. Educators are usually directing the accountability focus on to others and away from themselves. Conversely, from their experience of accountability systems generally in their lives the public would probably expect educators to have a directly responsible even if not solely responsible role in the accountability of the education system.

Factor 5 – Teacher incentives versus Respondent type

Simplistically, accountability may spur improvement by the linked use of incentives for desirable performance or/and disincentives for non-performance. Teachers use stars; high grades; and other accolades for desirable performance whilst society generally uses the withholding licences and permits; or fines and imprisonment for non-performance. It is not unexpected that 'incentives to teachers' is viewed differently by the two groups.

Factor 5 – Teacher incentives versus Gender

It is also not unexpected that there is correspondence between gender and how teacher incentives are viewed. Most teachers in Bermuda are female. Typically, females in western society are brought up to expect to be rewarded more often than they are punished as attempts at motivation and encouragement of appropriate behaviours. Males often have quite the converse experience and expectation.

Factor 10 – Reward student performance versus Where pre-college schooling took place

This difference might reflect two different features of the respondent population. The fact that pre-college schooling took place other than in Bermuda might indicate that the person is not a native of Bermuda (originating in another cultural and societal milieu) or comes from the upper socio-economic group in Bermuda who could afford to and did send their youngsters overseas to boarding school. The instrument may be modified to address this ambiguity of the socio-economic aspect as against differences due to geographic and ethnic origins. This might be a question on national origins or birthplace.

Bermuda is a fairly conservative society with a mixture of educational practices drawn from the United Kingdom, North America and the Caribbean that is superimposed on the Bermuda manner. Much of the English Caribbean has a more disciplinarian and strict style of public education. In North America it is more supportive. The public education system is not as class divisive as in the English grammar school versus comprehensive tussle but it is not as decentralized and laissez-faire as that of the USA. This focus on student reward might reflect experiences of these three locales.

Factor 6 – Preparation for work versus Highest level of education

This is a very logical interaction when one knows Bermuda well and even when one considers a common objective of education. Preparation for work, or more properly for early admission into what is usually less remunerative and also less intellectually demanding work, is often seen as an option for those who do not wish to seek or do not seem 'suited' for higher education. Generally the two are seen as in opposition. Those persons who are pursuing higher education do not expect to be entering the world of work very soon. It is also a common jibe that higher education does not prepare one well for the world of work.

Bermuda's twin economic pillars are tourism and servicing of international business. The former is labour intensive and not requiring of very high education for most positions. However, servicing of business is a globally competitive arena as modern telecommunications means that local service providers need not be used unless they offer significant advantages. International business uses the brightest and the best of Bermuda and elsewhere. The less educated more quickly prepare for and enter the work world.

The factor analysis opened the way for a far more searching analysis and indicated directions for further research. To some extent it might be possible to do further statistical analysis on the data already gathered and draw further inferences from the factor scores and ANOVA but this was considered unwise. The current data were just creditable for meeting criteria that were required for valid factor analysis. The sample size was just adequate as determined by the K-M-O (Kaiser-Meyer-Olkin) Measure of Sample adequacy and by Bartlett's Test of Sphericity (which assessed the patterning in the correlation matrix to determine its suitability). An increased sample size to at least 200 useable returns would be the first step.

Chapter 6 – Significance of the study

Introduction

This research project produced findings of value from two very different perspectives. The first finding is a project instrument that could assess the views of different stakeholder groups as regards an accountability system for K-12 public education. It is pointless to talk about accountability and accountability systems for education if we do not know what the stakeholders feel/think about the accountability system. We also want to know what it is that makes a system seem successful or less successful in the opinion of the respondent. We might expect that different respondent groups with different objectives and backgrounds would have different opinions on education accountability issues. The first finding of this research is that the methodology, although based on guidelines and theory and features with a North American genesis, was effective in a foreign context ---i.e. Bermuda. This suggests its broad applicability.

Secondly, these results are specific to Bermuda, as is typical of a case study, yet belong to a body of global data and contribute to the field of relevant literature concerning the accountability of accountability systems in pre-college public education.

It is important and useful to have a means to gather information about stakeholder opinions on public education. This is more informative than the system that is utilized by the Phi Delta Kappa/Gallup Poll "Of the Public's Attitudes Towards the Public Schools" (PDK/GP). The PDK/GP is a national overview of the public's opinions but this is a study at a local level. Local accountability for education as a local issue is a mark of modern activism. Therefore, the PDK/GP is used as a reference and comparison.

Value of the public's opinion on public education

Accountability has become life's universal watchword. It has assumed increasing importance for public education as the importance and accessibility of a 'good, basic education' often determines the success of the individual and the community (Vinovskis 1998; Black and Black 1987; Kozol 1991). In a democratic society, systems are managed

within the framework that the public consensus will support or tolerate. Thus, the managers of society must be able to measure the public's opinions and perceptions by a clear and accurate means. Education of the general public is fundamental in the egalitarian society that America strives to be (Tyack, 1999). Without effective and universal basic education there follows the divisive and caustic consequences of the sharp inequities of society noted by Kozol (1991). Yet public education in the USA has traditionally been the responsibility of local government rather than the federal system (Vinivskis, 1998; Black and Black, 1987; Tyack, 2000). In fact polls show that Americans "have trusted local school boards to make educational decisions far more than they trusted federal or state officials" (Tyack 2000). In the quest to reinforce local accountability there are new players and new strategies. Thus the 2001 edition of Standard and Poor's School Evaluation Services (SES) calls itself "a powerful decision-making tool for taxpayers, educators, school boards, and state and local education policy makers and administrators." It has academic, financial, and demographic indicators and trends; provides valuable comparative benchmarks; and presents Standard and Poor's "impartial findings, together with supporting data, on the performance of school systems." The Business Roundtable was founded in 1972 as a collective of CEO's of major USA corporations with the intent to have an increased role in debates about public policy matters. An Education Task Force would later focus specifically on "improving the performance of our schools, from kindergarten through grade 12, in each state" (Business Roundtable 1998). The Business Roundtable supports direct research as well as assists the publishing and dissemination of education information. These approaches demonstrate the present trend of an increased information gathering and attention to local public school systems (Finn,Jr., 1999; Tyack, 1999; Horsch, 1996; Manno, 1998).

PDK/GP has been a standard for long-term study of the public's opinion about its school systems. PDK's annual survey of public education perspectives ---"The Annual Phi Delta Kappa/ Gallup Poll Of the Public's Attitudes Towards the Public Schools (PDK/GP)" was in its 30th year of operation in 1998. The PDK/GP is a telephone survey of a random sample that is based on "a stratified and proportionate sample design; and is otherwise structured to provide results that can give great mathematical and statistical confidence in the validity of any findings" (PDK/GP 1998; Rose and Gallup.1999; 2000; 2001). The PDK/GP findings

"apply only to the U. S. as a whole and not to individual communities. Local surveys, using the same questions, can be conducted to determine how local areas compare with the national norm." (Rose and Gallup, 2001).

The PDK/GP data are broken down 'into 12 major population groups and 50 subgroups. "It is possible to compare the responses given by those living in urban, suburban and rural areas. Comparisons can be made of the responses provided by men and women or by Democrats, Republicans and Independents." (PDK/GP 1998). Since 1999, interview questions have been added to assess the respondent's educational background. Data have shown associations between the educational level of the respondent and a preference for student letter grades or student scores on standardized tests as an outcomes-based assessment method for schools. In the 2001 survey, data were sought on the income of the respondents and the size of the communities to examine these for any correlations.

We note that, "the original reason for initiating the Phi Delta Kappa/Gallup Poll was to provide those making decisions about the schools with the data to be used as input in those decisions." (PDK/GP 1998). Such data about opinions are important in making decisions in a democratic society. In the USA's decentralized education system, the general public must support a decision if it is to be easily and successfully implemented. Therefore, all education managers must first know the public's opinions on the issues (Business Roundtable 1998; 2000). It is disastrous to be ignorant of such opinions and to be faced with public apathy or lack of financial support or outright opposition.

Comparisons between the PDK/GP survey and the research project

In the 1998 to 2001 PDK/GP surveys there is consistent focus on accountability issues and strategies and issues of public school improvement with coverage of much that is also addressed by this research. Despite academic debate about the use of management tools in the educational setting (Reis, 2002; APQC 2002; Business Roundtable 1998; 2000) more persons daily encounter and use such tools for measuring performance as a first step in an accountability scheme (NEGP 1999; NEG Report 1998). This project uses such a tool to measure stakeholders' opinions about issues in public school accountability and education

systems as does the PD/GP survey. This research, however, secures more descriptive detail through respondent rated answers to a number of questions that focus on public school accountability. Some of these questions are also found in the PDK/GP survey while some are derived from rating systems in use in various educational settings or from a pilot survey of selected educators and students of UT-Austin.

Both the PDK/GP survey and this study have demographic information that permits the separation of the sample into various sub-groups for comparison as regards the response patterns. The actual number and distribution of responses is more important than the response rate or frequency of return. The former ensure validity while the response rate reflects the efficiency of the implementation process. The PDK/GP survey uses a sample of about 1100 persons aged 18 years and older from a national sample that is unclustered; random-digit telephone number assigned in a proportionate stratified sample design with weightings to match the data of the current Census Bureau Survey (PDK/GP 1998). This represents data for about 200 million persons in the USA as a whole (PDK/GP 1998; Rose and Gallup 1999; 2000; 2001). In this study a sample of 166 provides data for a population of about 55,000 and compares well with the ratio of sample size to population in the PDK/GP series and is statistically sufficient. The ratio used in this research would even be generous for the populations of many individual states by this comparison.

The PDK/GP surveys (2002.2001) also link sampling size and the recommended error allowance that should be used. Logically, smaller sample sizes show a greater chance of inconsistency in findings while larger samples should be more uniform and show less variability. The PDK/GP survey tables of sample size and recommended error allowance show that a sample size of 400 has only a slightly larger error allowance than a sample size of about 1000. A sample of 400 has slightly more inherent variability in the findings. If these same or similar sampling and recommended error allowance relationships hold for the methodology used in this project, then the research methodology used in this project can easily be applied to many systems in the USA. From the field experiences seen during this research it would not be difficult to obtain 400 to 600 returns from much larger populations. This involves about three to four times the number of responses in this case study (if we

assume a similar rate of return) but the populations of such regions are usually at least 20 to 30 times greater for individual state systems. This case study yields information on a methodology that has low resource requirements of personnel and materials and which is easily implemented but which can be at least as useful as a well-established and national survey such as the PDK/GP survey.

This survey was more informative than the PDK/GP in focusing primarily on opinions about accountability issues and improvement in the context of public K-12 schools. In this research project opinions of the different stakeholder groups were mapped further in some depth by exploring the means of the responses to the items considered, by using indexed scores and then by ANOVA studies of factor scores and the demographic variables. Divergences of opinions were more easily distinguished and perused more profoundly by these techniques. The study instrument was more focused and detailed than the PDK/GP tool but both suffered from the constraints of using standardized items.

Significance to Bermuda

Opinion results may be as particularly important to Bermuda (Saul 1973) as the PDK/GP survey findings are in the USA. In the Bermuda democracy, if perception is reality and opinion follows from perception then surely the opinions of the stakeholders are critical to the Bermuda education system (Saul 1973). In this research the stakeholders gave their opinions about the relative importance of a variety of selected factors associated with accountability in the context of public K-12 education in Bermuda.

These results will need specifically to be viewed against the backdrop of features that are peculiar to Bermuda. Further analysis of these results must be framed to include consideration of material such as the Annual report of the Ministry of Education of Bermuda; the Annual Report of the Board of Education; the Educational Planning Team analysis (1989); The Middle School Review (1994) and the Curriculum Management Audit (1996).

This study may eventually contribute directly to the relevant research about the operation of the education system and about Bermuda in general. Bermuda lies between the USA and the

UK and draws heavily from both for some staffing, management and operation needs. The importance of this study is related to the specificity of the results to Bermuda and the unique aspects of the Bermudian educational system as placed in the galaxy of things Bermudian and which affect Bermuda. These will include local political changes; local economic changes; local lifestyle changes and the role that the education system plays in current and future status of each of these parameters.

Significance to USA jurisdictions

But the case study aspect means that this research has significance outside of its value within the Bermuda sphere. Other than the PDK/GP Annual Survey of the Public's Attitudes towards Public Schools, there has traditionally not been a large body of research about stakeholder opinions (Roueche, 1997; 2001; National Center for Education Statistics, 2000; Ruppert, 1998). The concept that stakeholder opinions matter is relatively new in the broad life of society and still is not universally accepted in education circles. This joins mid-twentieth century concepts of not simply pleasing the customer but of anticipating the customer's needs and desires and then meeting and surpassing them (APQC 1999; U.S. Department of Education 1993; Deming 1993). This calls for constant monitoring of the actual situation as regards products, processes and services as well as the customers opinions about these same factors so as to produce a never-ending stream of improvements. This was the mantra of Continuous Quality Improvement (CQI) and Total Quality Management (Drucker, 1999; Juran, 1993).

But as education has come to assume ever-growing importance in the success of the individual and the nation it has become clearer that some assurance of the soundness of educational preparation has become necessary (Finn, Jr. and Petrilli 2000; Rahn and Holmes 1999). Accreditation agencies like the Southern Association of Colleges and Schools (SACS) provide minimal standards that are to be reached as some sort of indicator of the base quality of the preparation process. These minimal standards for various parameters provide the starting point for some accountability in the school system's educational performance as regards the preparation process for the student. Such indicators are shown in publicly available profiles of school systems from the National Association of School Boards

of Education (NASBE) or the Council of Chief State School Officers (CCSSO). In the present climate of anxious debate about the effectiveness of K-12 public education (Rahn and Holmes, 1999; Wellman, 2001), reports are also obtainable from CCSSO on the status of the accountability system for education in each state with an overview of the schools in the state (CCSSO, 1999).

Most published information about public K-12 accountability systems is descriptive rather than analytical (CCSSO, 1999; Standard and Poor's School Evaluation Series, 2001). Progress among the various states has been so haphazard that accountability systems vary greatly in development (Finn, Jr. and Petrillili, 2000). Despite the efforts of the National Education Goals Panel there has been limited progression to directly comparable accountability systems or education systems. Consequently stakeholders are faced with a bewildering array when attempting to develop an opinion on the efficacy and validity of these accountability systems in education (Public School Accountability Act 2000; RAND Research Brief 1999). But such an opinion needs to be measured.

The importance of the validity and efficacy of the accountability system in education is not to be belittled. As the world status of the USA is dependent on national performance in sports and in artistic endeavours and has become crucial to its self-image and its prestige among other nations, so too has its educational ranking become very important (International Education Indicators, 2000). Without a consistent accountability process the various state officials often did not even collect and track data on the indicators that were required for international ranking. At other times the data were incomplete or so inconsistent when gathered from the myriad of systems that exist in the different and autonomous units of the US K-12 public system that no clear or total picture could be drawn from them. Over a little more than the past decade the NAEP and the NEGP have led the way in promoting the measurement of educational progress with the intent of having consequences and responses to promote its advance.

On the one side is the historical distinctness and autonomy of state education systems with each local population being given authority to develop a system that it judges to be

appropriate to its needs with freedom from prescription and control by any federal authority. On the other hand is the current focus on accountability to ensure that present and future systems meet the needs of all citizens in an equitable and fair manner. The perception that there would be stringent accountability has been shown recently to be associated with dramatic and remarkable improvements in student performance. Such improvement is the objective of all accountability systems. In the jargon of the TQM and CQI proponents this reflects the consequences of managing so as to maximize performance of the inputs (the efforts of teachers and administrators and parents, etc. as well as the physical aids etc.) and reaping an improved output. This is outcomes-based accountability. The idea of the application of such accountability techniques may appear radical in the USA but is the norm in many other parts of the world in Europe, Australia and Asia (Dill, 2000; Trombley, 2000).

The Texas education accountability system has achieved pre-eminence for a number of reasons. Texas is one of the largest states and has one of the largest public K-12 systems in the USA. Indeed it, along with the huge California system, acts as the standard for textbook adoption and use of various curriculum aids. Texas was also one of the first locales to have a state accountability system developed and this system is one of the best known and understood nationally (Finn, Jr. and Petrillili, 2000). Despite some criticisms the Texas system has received high praise overall (Jerald, 2001;Texas Education Agency 1998; Haney, 2001; CBE, 1998). More attention has been garnered with the accession of the former governor of the state to the US presidency as President G. W. Bush and with the former Dr. Ron Paige of Houston becoming the national Secretary for Education.

There is a startling confluence of factors as regards the issue of accountability systems in education and the education system of the state of Texas. The Texas system is one of the largest; has one of the longest operating and most stable accountability systems in the USA; and has been shown to follow the predictions of CQI and TQM adherents about the ability of stringent accountability to drive performance to unexpected heights of quality and success when stringent accountability is perceived as a basic component of performance (Scheurich et al, 2001). Certainly the application of the PDK/GP survey at a local state level would be informative but it is also clear that much more information could be provided by using the

format employed in this study. The first step would be a questionnaire survey based on the exact items used here. Additional demographic questions would be added for ethnicity and such data. The major question section, with the items on accountability and related issues in an educational context, could be used probably without major changes.

Interestingly, tertiary educators are also eagerly awaiting the developments arising from such an analysis of issues of accountability in education systems for K-12 public education. This pertains to the publicly funded community and state colleges and universities as well as the private institutions (Mize, 199; Ruppert, 1998; Alfred, Ewell et al 2000, Community College Roundtable 1991, Wellman, 2001; Ewell, 2001). Tertiary education is increasingly necessary for individual and national success in an ever more complex world. Yet tertiary education becomes more expensive and more difficult to support/obtain as the numbers of persons, who want to take advantage of it, explode exponentially (Callan et al, 2000). This shows the need for greater efficacy and efficiency in the publicly funded education process. In the non-education world such heightened improvement and accountability have been achieved by the adoption of adapted principles of CQI and TQM (APQC 2000a; 2000b; Business Roundtable 1996). It seems to many that a like situation must follow in education (Ohio Board of Regents, 1995; Wells, 1999; Alexander, 2000). Modified forms of CQI and TQM will need to be adopted. A first step in TQM and CQI is accurate measurement and then monitoring of customers' perceptions of performance as well as measurement and tracking of the real indicators of performance. This research project provides an option that, by the case study, is established as viable. It can be applied for tertiary education as it has been and is intended for the public K-12 system.

Conclusion

In summary, the research provides a model for further inquiry about issues in accountability and improvement for the public K-12 school system in Texas and elsewhere in the USA. The Texas public school system has had particular notoriety as regards issues of accountability in education because of a history of distinctly separated educational opportunities for its various ethnic populations (Black and Black, 1987) and its rapid move to a more equitable system of education through the implementation and successful operation of

accountability systems (TEA, 2000). Texas has become the capstone state for accountability in public school systems and is seen to be at the leading edge (Finn and Petrillili, 2000: Business Roundtable, 2001). Of extreme interest has been a set of very recent findings that suggest strongly that stakeholders' opinions and perceptions as they pertain to accountability issues can in some instances at least lead to dramatic improvements in school performances. This improvement is measured by increases in the outcomes-based assessments of student performances (Scheurich, et al 2001; Scheurich, et al 2000; Johnson, 1998).

It is believed that, with some minor adjustments to the questionnaire perhaps being necessary, the general format would also be transferable to tertiary education analysis where issues of accountability are important because of limits of access; the heightened cost of tuition and other items and the longer established and stronger tradition of non-accountability to stakeholders that exists in tertiary education in the USA. This new focus on the customer in education has greater concern because higher education is considered less of a public responsibility than is the public K-12 system (Ruppert, 1998; Alexander, 2000; Callan et al, 2000).

This research project may be a goldmine of data about the opinions of various stakeholders of Bermuda on issues of accountability in public K-12 education. Bermuda has had more externally generated analyses than those that have been conducted with a local perspective (Saul, 1973; Curriculum Management Audit 1996; Bermuda College External Analysis and Audit 2000). Much data are collected for the numerous local reports but no data are routinely collected about the public's attitudes or opinions towards the public school system. Recent comments in the legislative settings (Parliamentary Debate on the Education Budget 2002) and public opinion as expressed on the street would suggest a probable dissatisfaction with the performance of the public system. This research offers a model of how the details of the public's opinions could be clearly and confidently secured. By analyzing the specific details of associated issues in the accountability of the system in Bermuda and how these items vary for stakeholders, it will enable the managers of the system to be better informed as to how to respond in the most appropriate manner. Sometimes this may be a requirement to provide information to correct a negative perception. At other times the analysis of the

survey items which frame the negative opinion will reveal what means need to be addressed to allow that opinion to be reversed. Conversely, items that denote a strongly positive opinion will describe areas of success and where focus can be maintained as necessary.

Appendix A

Format of questionnaire:

Type 1 questions are questions 1 to 5 (or A to G).
These inform about the demographic features of the responding sample.

Type 2 questions are questions 6 to 15 (or H to R).
These supply contextual information of the respondent's opinion on values and objectives of education and are influential in determining their views on accountability issues.

Type 3 questions are questions 16 to 71 (or S to BV).
These supply the data of the respondent's opinions on issues and associated items of outcomes- based accountability systems. Questions and items are derived from performance indicators used in accountability rating systems and from reflections on the theory of outcomes-based accountability and quality assurance. Most items were pre-tested in a small pilot among education students at UT-Austin.

Themes and the questions associated with them in the survey are noted below.

School administration	17, 34, 35, 45.
School curricula	16, 18, 22, 23, 25, 28, 30,.31, 33, 36, 43, 44.
School behaviours	20, 21, 24, 29, 31, 37.
Personal inputs/rewards	26, 32, 41.
Non-school system factors	27, 35, 38, 39, 42.
Direct link to OB accountability	16, 34, 41, 45, 49, 51, 59, 62, 68.

A sample questionnaire follows.

This is a 3-page document that is presented exactly as seen by the respondent. There is a cover page and/or a verbal explanation that also are given. This is very brief and simply introduces the researcher and the purpose of the research. The sample questionnaire follows overleaf.

GENERAL INFORMATION

Please answer all questions. This information will assist in analysing the results for patterns. Shade the number in brackets that corresponds to your answer.

1. Age (1) under 20 (2) 20-24 (3) 25-29 (4) 30-49 (5) 50+

2. Gender (1) male (2) female

3. Pre-college schooling (1)Bermuda (2) USA (3) UK (4) Other

4. Occupation (*please print*) _____

5.

Highest level of education	**(1) Primary school**	**(2) Secondary school**	**(3) Bachelors**	**(4) Masters or Doctorate**	**(5) Professional Training**
Time of last educational event	**(1) currently going on**	**(2) within the past year**	**(3) within the past 2 years**	**(4) within the past 5 years**	**(5) more than 5 years ago**

PURPOSES OF EDUCATION

*Indicate how important you think it is that the public school system should be excellent at the following features by circling a number. **1** shows least importance. **5** shows highest. **0** means not relevant.*

6. Preparation for college/university	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
7. Teach responsibilities of citizenship	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
8. Preparation for work/profession	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
9. Impart social values, morals etc.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
10. Develop individual areas e.g. confidence	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
11. Sustain community culture, traditions, etc.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
12. Teach life skills	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
13. Develop appreciation of art, learning etc.	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**

14. Name any 2 other purposes of education and attach a number between 1 and 5 to show their importance.	* *
15. If only 3 of these purposes could be carried out by the public school system, which should they be. (In order of highest importance first)	* * *

Rating School Attributes

*For the following, **1** shows least importance, **5** shows greatest. **0** means not relevant.*

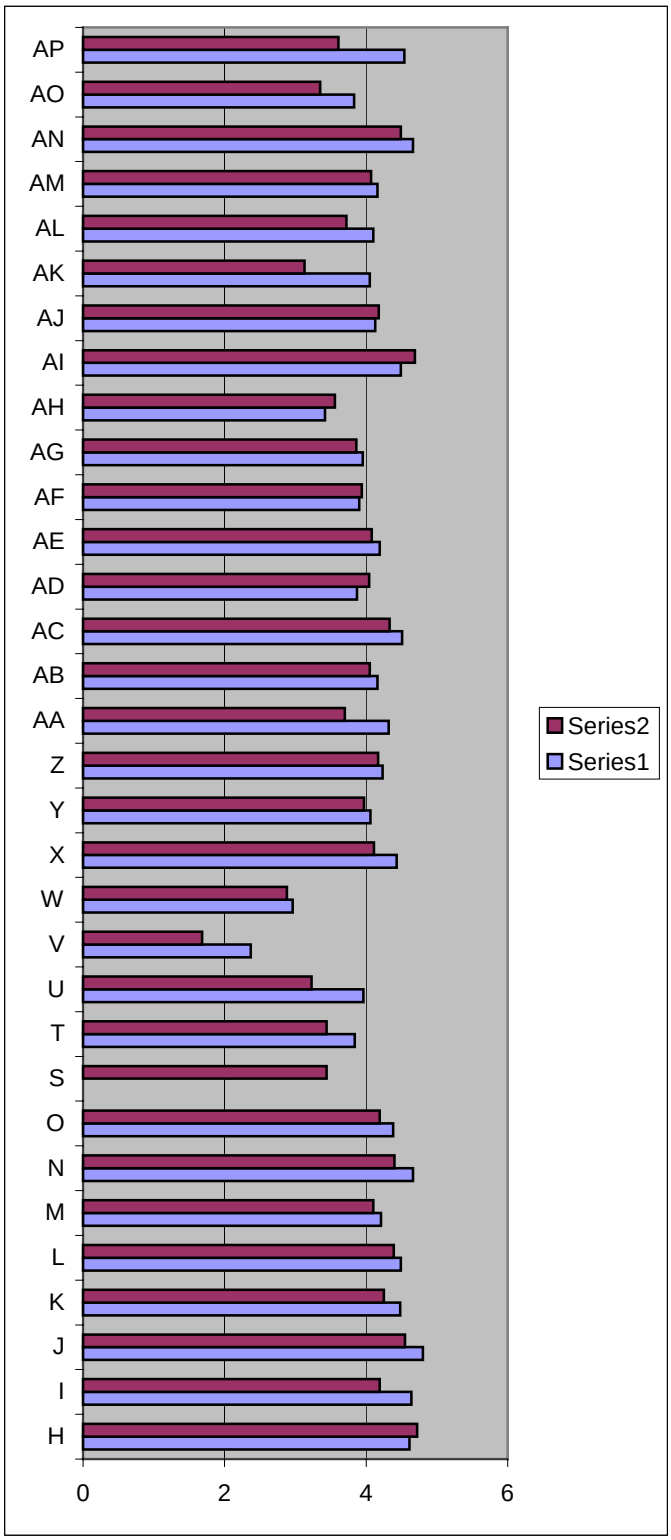
16. Schools have publicly published ratings	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
17. Multi cultural awareness classes	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
18. Community service needed for graduation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
19. School may forbid part time student jobs	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
20. Disabled students are main stream classes	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
21. Extra curricular activities (band, sports)	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
22. Basic technical subjects are compulsory	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
23. Learning has team work components	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
24. Music, art, poetry etc for all students	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
25. Teacher's empathy is most critical quality	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
26. Career guidance should include parents	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
27. Everyone learns basic economics/business	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
28. People skills are a core requirement	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
29. Trade/job internships for graduation	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
30. Ethics and morality are required topics	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
31. Daily homework including weekends	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
32. Teachers are certified for their subject	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
33. Maths requirements should be raised	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
34. Teachers/principals get incentive awards	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
35. Persons on school boards receive training	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
36. Civics, national history etc are compulsory	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
37. Students get extra lessons in deficient areas	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
38. Family members volunteer at school	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
39. Parental school participation is mandatory	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
40. Social promotion is abandoned	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
41. Remedial class teachers get success bonuses	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
42. College/school collaboration is mandatory	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
43. More problem solving strategies are taught	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**
44. All learn basics of computer studies	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(0)**

For the following, show your level of agreement. ***1*** *is least strong and* ***5*** *is strongest.* ***0*** *means not relevant.*

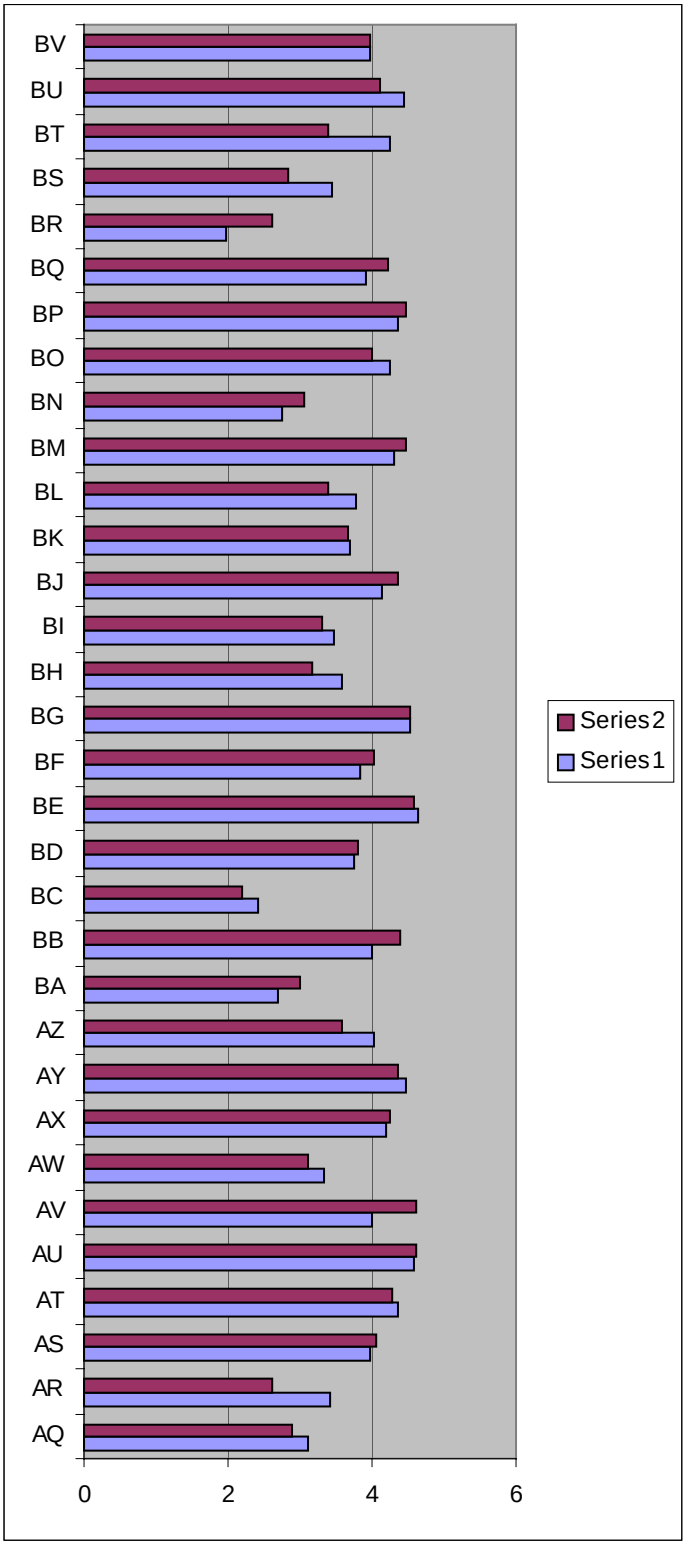
45. Principals should be more accountable for the schools. (1) (2) (3) (4) (5) (0)
46. Students should have more say in school accountability. (1) (2) (3) (4) (5) (0)
47. Smaller class size should be accompanied by better results. (1) (2) (3) (4) (5) (0)
48. Parental attitude to school determines student motivation. (1) (2) (3) (4) (5) (0)
49. School rating must not be overly dependent on test results. (1) (2) (3) (4) (5) (0)
50. We should consider extending the school year for students. (1) (2) (3) (4) (5) (0)
51. Teachers should be able to show the quality of their work. (1) (2) (3) (4) (5) (0)
52. Schools meet student's emotional needs better than parents. (1) (2) (3) (4) (5) (0)
53. Students are not working sufficiently hard at school. (1) (2) (3) (4) (5) (0)
54. Early parental reading to students is a family duty. (1) (2) (3) (4) (5) (0)
55. Schools must teach more work skills like timeliness. (1) (2) (3) (4) (5) (0)
56. Teachers must have time to keep skills updated. (1) (2) (3) (4) (5) (0)
57. Communities must supplement government school funds. (1) (2) (3) (4) (5) (0)
58. Parental education determines children's school attitude. (1) (2) (3) (4) (5) (0)
59. Unsuccessful teachers should be retrained or removed. (1) (2) (3) (4) (5) (0)
60. Ability to motivate is more important than credentials. (1) (2) (3) (4) (5) (0)
61. Teachers determine success indicators for education best. (1) (2) (3) (4) (5) (0)
62. Education should be improvable just like other professions. (1) (2) (3) (4) (5) (0)
63. Bad teaching does more harm than poor family structure. (1) (2) (3) (4) (5) (0)
64. Business must invest more in schools for their own futures. (1) (2) (3) (4) (5) (0)
65. Education success yields both social stability and progress. (1) (2) (3) (4) (5) (0)
66. Higher education must assist in upgrading public education. (1) (2) (3) (4) (5) (0)
67. Private education is always better than public education. (1) (2) (3) (4) (5) (0)
68. Public education needs to be operated more like a business. (1) (2) (3) (4) (5) (0)
69. Teachers need better compensation for their academic level. (1) (2) (3) (4) (5) (0)
70. Society's attitude to education needs a lot of improvement. (1) (2) (3) (4) (5) (0)
71. Society's negativity about school is mimicked by children. (1) (2) (3) (4) (5) (0)

Appendix B

1.0 – Bar charts showing means of matched question responses by general public (upper bar) and educators (lower bar).



1.1 – Bar charts showing means of matched responses by general public (upper bar) and educators (lower bar).



.

Appendix C

1.0 – Descriptive statistics for questionnaire responses from the survey of educators. The top row identifies the questions.

	H	I	J	K	L	M	N	O
N	80	80	80	80	80	80	80	80
Mean	4.6125	4.6375	4.8000	4.4750	4.4875	4.2125	4.6625	4.3750
Median	5.0000	5.0000	5.0000	5.0000	5.0000	4.0000	5.0000	5.0000
Mode	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Std. Deviation	.8929	.6980	.6038	.7952	.7630	.8958	.6353	.9192
Variance	.7973	.4872	.3646	.6323	.5821	.8024	.4036	.8449

	S	T	U	V	W	X	Y	Z
N	79	79	80	78	75	79	80	79
Mean	3.0633	3.8354	3.9625	2.3718	2.9600	4.4304	4.0625	4.2278
Median	3.0000	4.0000	4.0000	2.0000	3.0000	5.0000	4.0000	5.0000
Mode	3.00	3.00	5.00	1.00	3.00	5.00	5.00	5.00
Std. Deviation	1.2944	.9665	1.1412	1.4781	1.2018	.7956	1.0106	1.0494
Variance	1.6754	.9341	1.3024	2.1846	1.4443	.6329	1.0214	1.1013

	AA	AB	AC	AD	AE	AF	AG	AH
N	79	77	80	79	79	80	79	79
Mean	4.3165	4.1558	4.5125	3.8734	4.1899	3.9000	3.9494	3.4177
Median	5.0000	4.0000	5.0000	4.0000	4.0000	4.0000	4.0000	3.0000
Mode	5.00	4.00	5.00	4.00	4.00	5.00	4.00	3.00
Std. Deviation	.9547	.9042	.7462	1.0423	.8332	1.1092	.9323	1.2154
Variance	.9114	.8175	.5568	1.0863	.6943	1.2304	.8692	1.4771

	AI	AJ	AK	AL	AM	AN	AO	AP
N	79	80	77	79	79	79	80	80
Mean	5.0000	4.1250	4.0519	4.1013	4.1646	4.6582	3.8250	4.5375
Median	5.0000	4.0000	5.0000	4.0000	4.0000	5.0000	4.0000	5.0000
Mode	5.00	4.00	5.00	5.00	5.00	5.00	5.00	5.00
Std. Deviation	1.0606	.9727	1.2555	1.1502	.9797	.6178	1.1449	3.5039
Variance	1.1250	.9462	1.5762	1.3229	.9598	.3817	1.3108	12.2771

	AQ	AR	AS	AT	AU	AV	AW	AX
N	71	74	79	80	80	79	80	79
	10	7	2	1	1	2	1	2
Mean	3.1127	3.4324	3.9747	4.3625	4.5875	4.0127	3.3375	4.1899
Median	3.0000	3.0000	4.0000	5.0000	5.0000	4.0000	3.0000	5.0000
Mode	4.00	3.00	4.00	5.00	5.00	5.00	3.00	5.00
Std. Deviation	1.6349	1.3455	1.0250	.8894	.8523	1.2454	1.2523	1.0387
Variance	2.6728	1.8104	1.0506	.7910	.7264	1.5511	1.5682	1.0789

	AY	AZ	BA	BB	BC	BD	BE	BF
N	80	79	77	79	79	80	80	77
	1	2	4	2	2	1	1	4
Mean	4.4750	4.0380	2.6883	4.0000	2.4304	3.7375	4.6250	3.8182
Median	5.0000	4.0000	3.0000	4.0000	3.0000	4.0000	5.0000	4.0000
Mode	5.00	5.00	1.00	5.00	3.00	5.00	5.00	4.00
Std. Deviation	.7952	1.1596	1.5666	1.1658	1.3838	1.2503	.8476	1.0225
Variance	.6323	1.3447	2.4542	1.3590	1.9150	1.5631	.7184	1.0455

	BG	BH	BI	BJ	BK	BL	BM	BN
N	78	80	79	78	80	77	80	78
	3	1	2	3	1	4	1	3
Mean	4.5385	3.5875	3.4684	4.1410	3.6875	3.7662	4.3000	2.7436
Median	5.0000	4.0000	4.0000	5.0000	4.0000	4.0000	5.0000	3.0000
Mode	5.00	5.00	5.00	5.00	3.00	4.00	5.00	3.00
Std. Deviation	.9493	1.4204	1.4308	1.2031	1.1648	1.0500	1.0721	1.6071
Variance	.9011	2.0176	2.0471	1.4474	1.3568	1.1025	1.1494	2.5828

	BO	BP	BQ	BR	BS	BT	BU	BV
N	78	79	78	79	78	80	80	79
	3	2	3	2	3	1	1	2
Mean	4.2436	4.3671	3.9103	1.9747	3.4487	4.2500	4.4375	3.9620
Median	5.0000	5.0000	4.0000	1.0000	4.0000	5.0000	5.0000	4.0000
Mode	5.00	5.00	5.00	1.00	5.00	5.00	5.00	5.00
Std. Deviation	1.0468	.9764	1.1863	1.3957	1.4201	1.1081	1.0415	1.1815
Variance	1.0957	.9533	1.4074	1.9481	2.0168	1.2278	1.0847	1.3960

1.1 – Descriptive statistics for questionnaire responses from the survey of the general public. N is the number of respondents. Top row identifies the questions.

	H	I	J	K	L	M	N	O
N	85	84	83	85	85	82	86	84
Mean	4.7176	4.1905	4.5542	4.2471	4.3882	4.0976	4.3953	4.1905
Median	5.0000	5.0000	5.0000	5.0000	5.0000	4.0000	5.0000	4.0000
Mode	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Std. Deviation	.7173	1.1137	.9140	1.0107	.8465	.9637	.7712	.9503
Variance	.5146	1.2404	.8354	1.0216	.7165	.9286	.5948	.9030

	S	T	U	V	W	X	Y	Z
N	84	85	84	84	75	85	86	86
Mean	3.2381	3.4353	3.2262	1.6786	2.8800	4.1059	3.9651	4.1744
Median	3.0000	4.0000	4.0000	1.0000	3.0000	4.0000	4.0000	4.0000
Mode	3.00	3.00	4.00	1.00	3.00	5.00	5.00	5.00
Std. Deviation	1.5100	1.3491	1.5550	1.4821	1.4516	.9639	.9994	.8567
Variance	2.2800	1.8202	2.4181	2.1966	2.1070	.9291	.9988	.7339

	AA	AB	AC	AD	AE	AF	AG	AH
N	84	81	86	83	85	85	85	84
Mean	3.7024	4.0494	4.3256	4.0361	4.0824	3.9412	3.8588	3.5595
Median	4.0000	4.0000	5.0000	4.0000	4.0000	4.0000	4.0000	4.0000
Mode	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Std. Deviation	1.3060	1.1169	1.1002	1.0054	1.0025	1.2850	1.1036	1.3828
Variance	1.7055	1.2475	1.2104	1.0109	1.0050	1.6513	1.2179	1.9121

	AI	AJ	AK	AL	AM	AN	AO	AP
N	84	83	83	85	85	86	84	83
Mean	4.6905	4.1807	3.1325	3.7176	4.0706	4.4884	3.3452	3.6145
Median	5.0000	4.0000	3.0000	4.0000	4.0000	5.0000	4.0000	4.0000
Mode	5.00	5.00	3.00	5.00	5.00	5.00	4.00	5.00
Std. Deviation	.6205	.8991	1.5832	1.3328	1.0211	.7315	1.5010	1.3689
Variance	.3850	.8084	2.5066	1.7765	1.0426	.5352	2.2529	1.8739

	AQ	AR	AS	AT	AU	AV	AW	AX
N	77	82	81	85	85	84	84	85
Mean	2.8961	2.6098	4.0494	4.2941	4.6118	4.5952	3.1071	4.2588
Median	3.0000	3.0000	4.0000	5.0000	5.0000	5.0000	3.0000	5.0000
Mode	5.00	3.00	5.00	5.00	5.00	5.00	3.00	5.00
Std. Deviation	1.7516	1.5851	1.2031	1.0098	.6380	.6787	1.4728	1.1355
Variance	3.0680	2.5125	1.4475	1.0196	.4070	.4607	2.1691	1.2894

	AY	AZ	BA	BB	BC	BD	BE	BF
N	86	84	85	83	84	84	85	83
Mean	4.3488	3.5833	3.0118	4.3976	2.1905	3.8095	4.5882	4.0361
Median	5.0000	4.0000	3.0000	5.0000	2.0000	4.0000	5.0000	4.0000
Mode	5.00	5.00	4.00	5.00	1.00	5.00	5.00	5.00
Std. Deviation	.9301	1.4075	1.5923	.8545	1.3751	1.1971	.8632	1.0412
Variance	.8651	1.9809	2.5356	.7302	1.8910	1.4332	.7451	1.0840

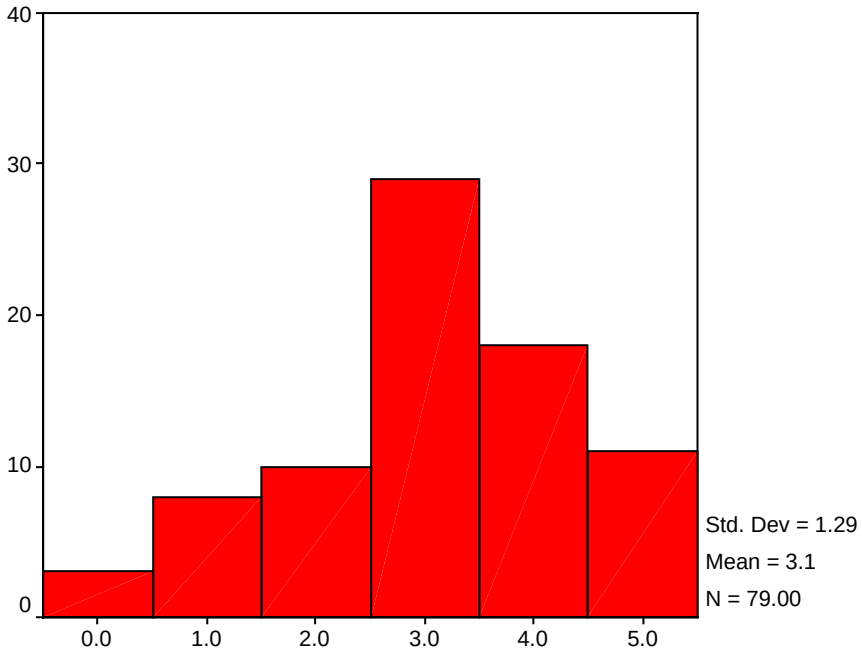
	BG	BH	BI	BJ	BK	BL	BM	BN
N	83	83	84	84	85	80	83	84
Mean	4.5301	3.1807	3.3095	4.3571	3.6588	3.3750	4.4819	3.0595
Median	5.0000	3.0000	3.5000	5.0000	4.0000	3.0000	5.0000	3.0000
Mode	5.00	3.00	5.00	5.00	5.00	3.00	5.00	5.00
Std. Deviation	.7214	1.4910	1.6055	1.2088	1.3052	1.2465	.9417	1.5704
Variance	.5204	2.2230	2.5777	1.4613	1.7036	1.5538	.8869	2.4663

	BO	BP	BQ	BR	BS	BT	BU	BV
N	84	83	83	84	82	84	83	83
Mean	3.9881	4.4819	4.2289	2.6071	2.8293	3.3929	4.0964	3.9759
Median	5.0000	5.0000	5.0000	3.0000	3.0000	4.0000	4.0000	4.0000
Mode	5.00	5.00	5.00	1.00	3.00	4.00	5.00	5.00
Std. Deviation	1.4013	.7865	1.1189	1.5756	1.6163	1.3799	1.0889	1.1891
Variance	1.9637	.6186	1.2518	2.4824	2.6125	1.9040	1.1857	1.4140

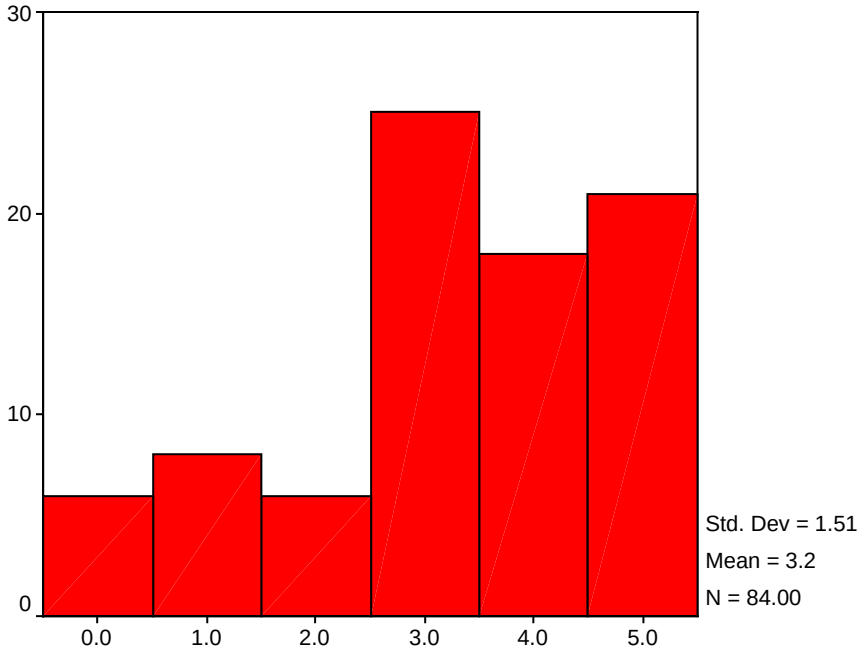
Appendix D.

Sets of matched histograms of the frequency of response categories

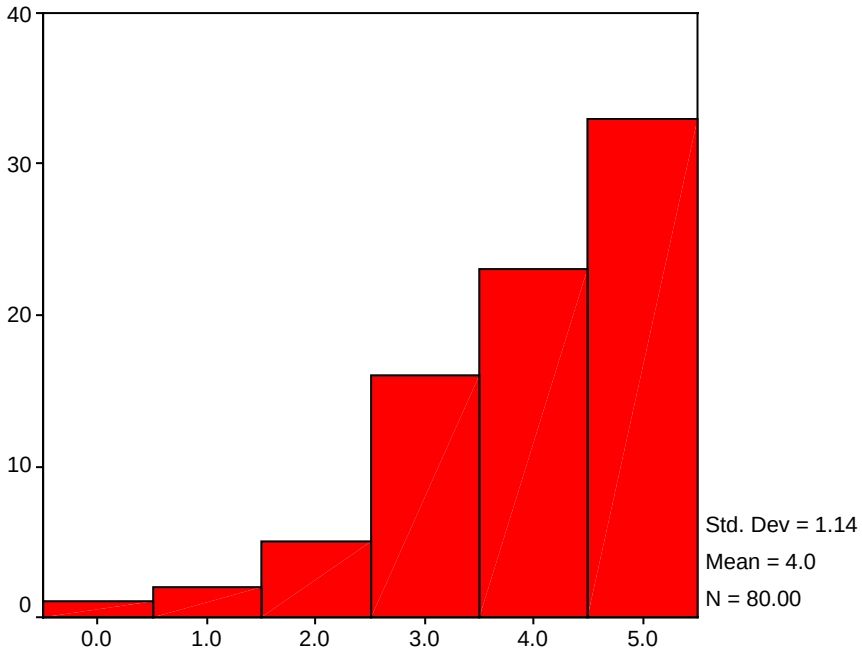
The relationships between these curves are described on p.76 of the text. In the sets of matched histograms the figure that depicts data from the educator sample is given in the upper graph and that for the general public sample is given in the lower figure.


40
30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.29
Mean = 3.1
N = 79.00

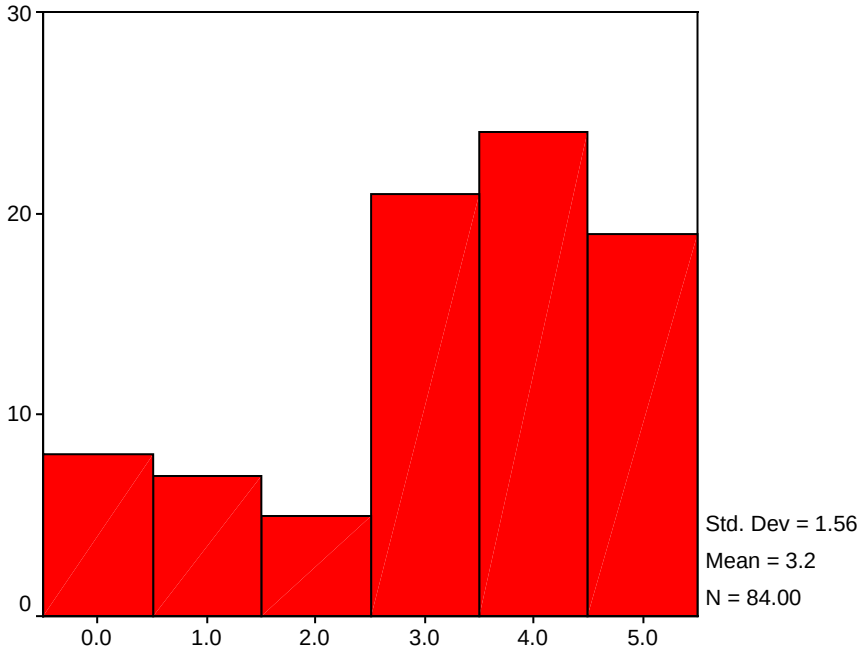
S


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.51
Mean = 3.2
N = 84.00

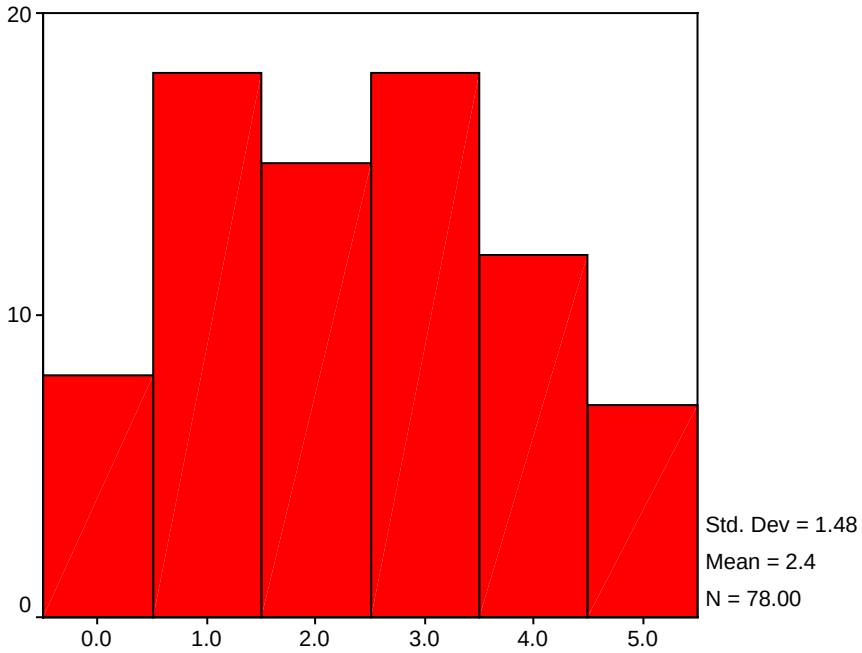
S

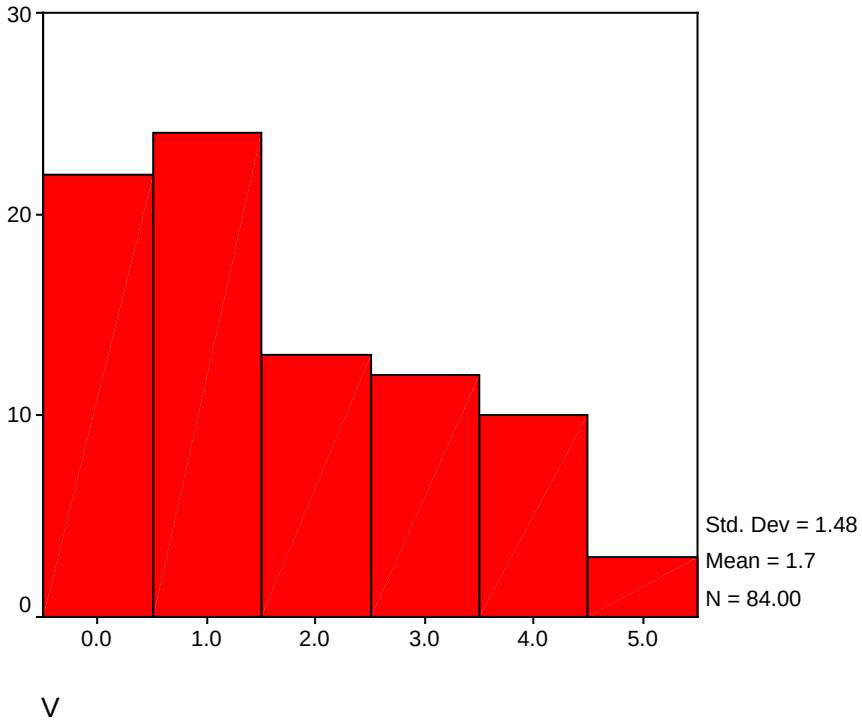

40
30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.14
Mean = 4.0
N = 80.00

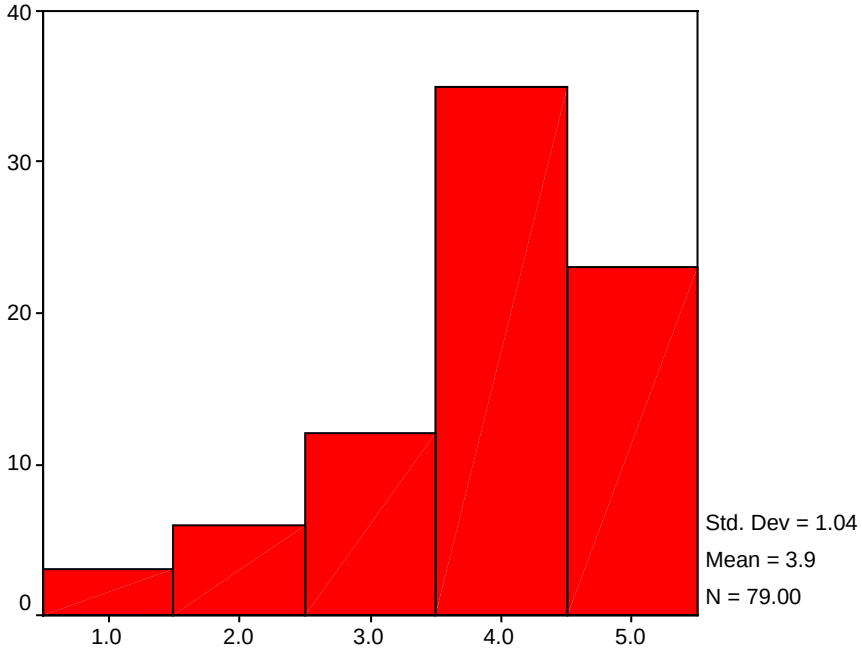
U


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.56
Mean = 3.2
N = 84.00

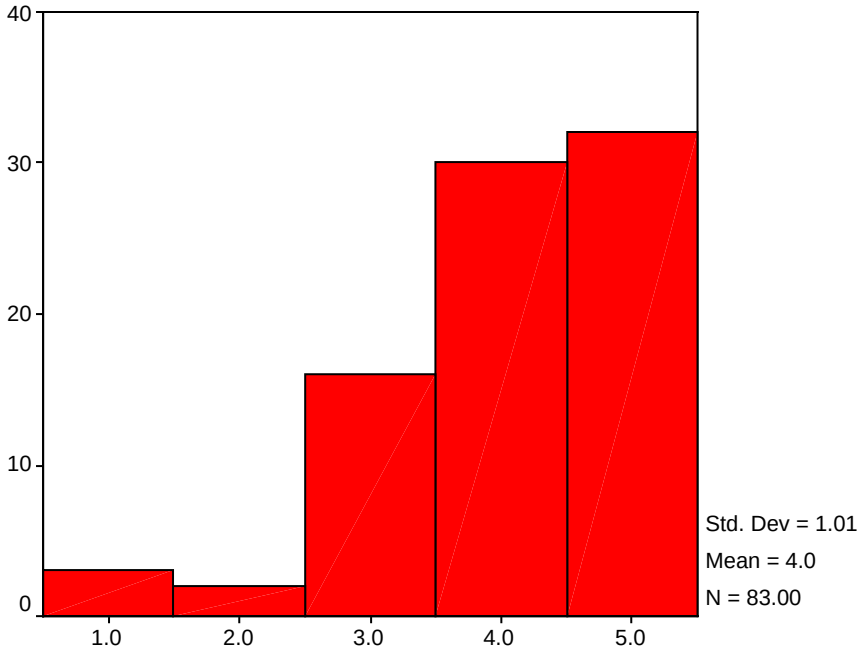
U


20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.48
Mean = 2.4
N = 78.00
V

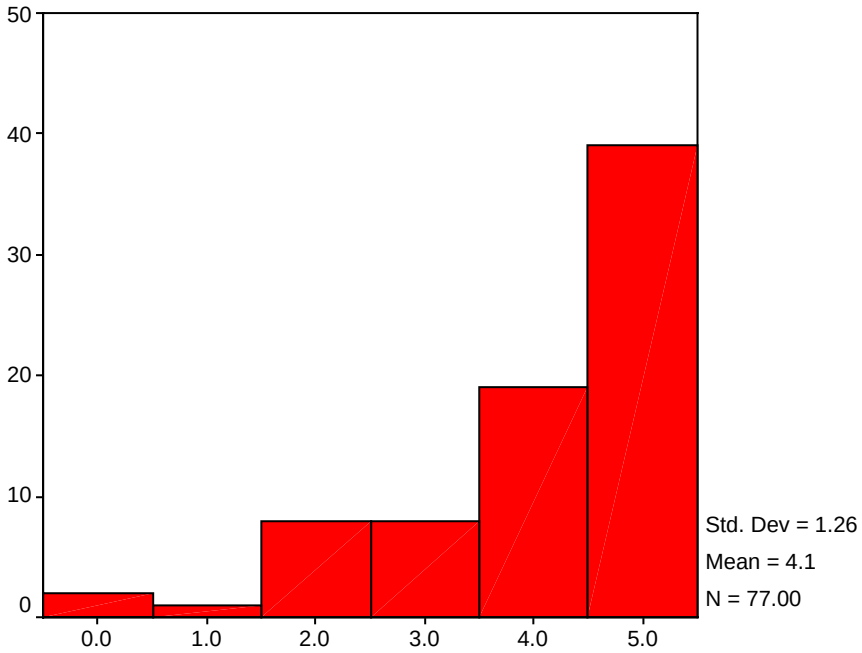

30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.48
Mean = 1.7
N = 84.00
V


40
30
20
10
0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.04
Mean = 3.9
N = 79.00

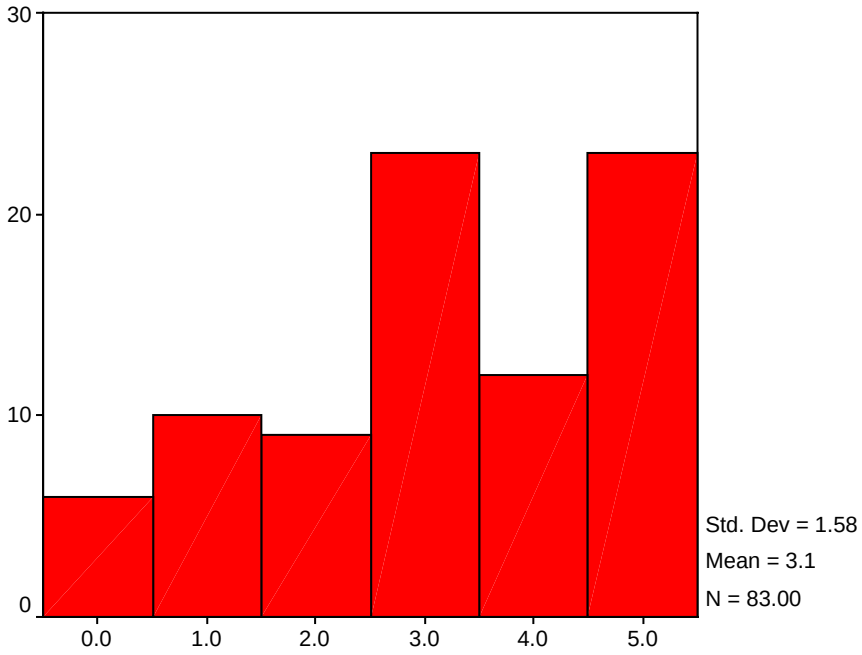
AD


40
30
20
10
0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.01
Mean = 4.0
N = 83.00

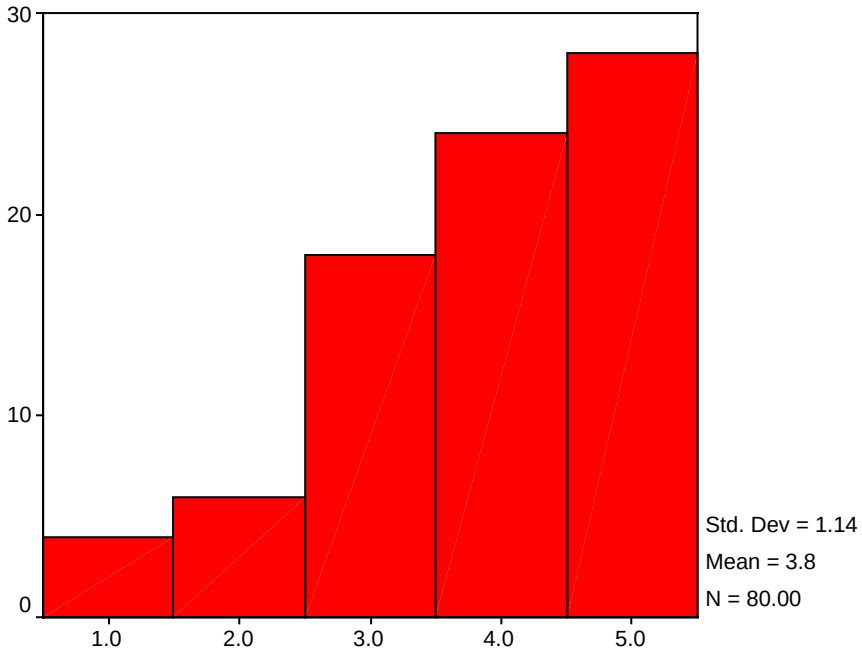
AD


50
40
30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.26
Mean = 4.1
N = 77.00

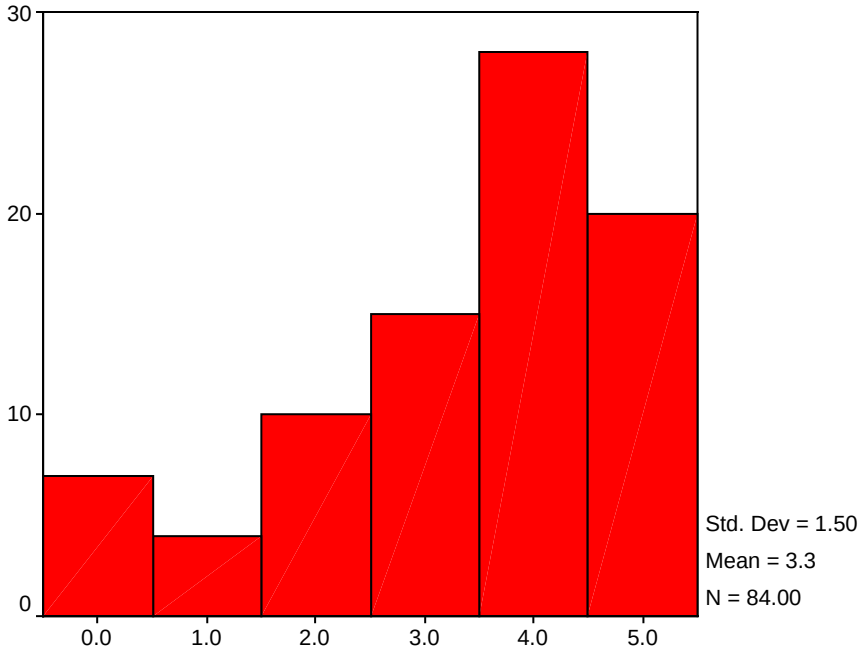
AK


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.58
Mean = 3.1
N = 83.00

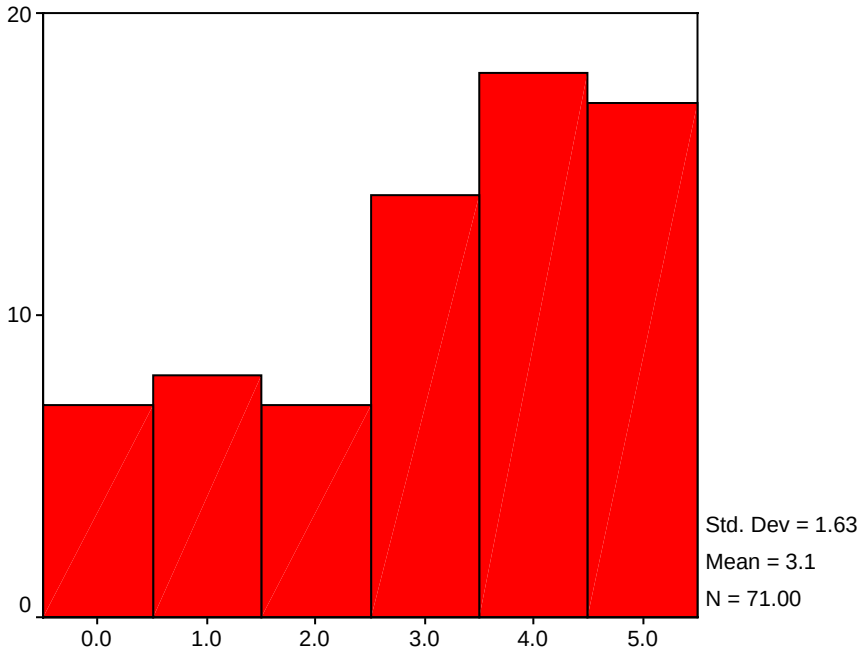
AK


30
20
10
0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.14
Mean = 3.8
N = 80.00

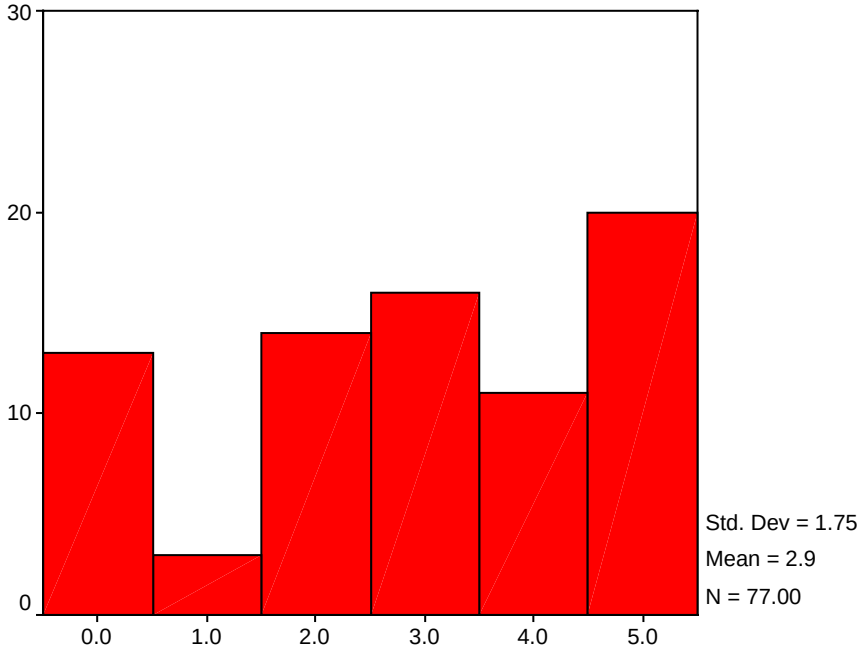
AO


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.50
Mean = 3.3
N = 84.00

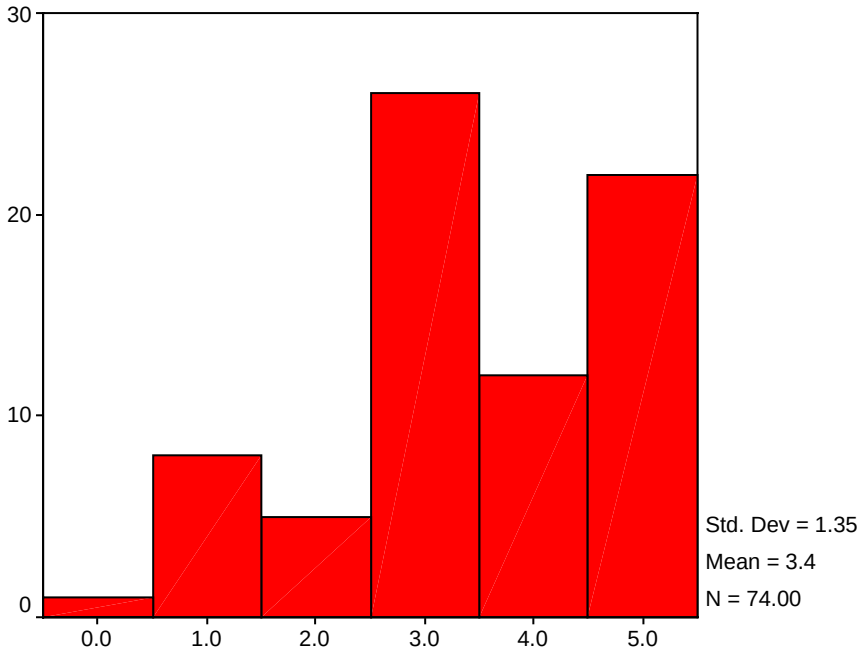
AO


20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.63
Mean = 3.1
N = 71.00

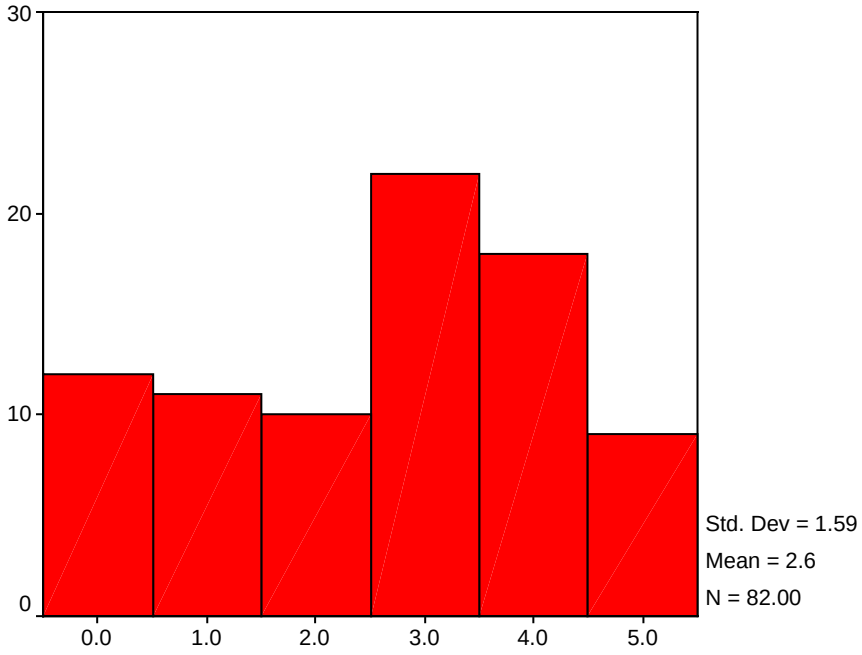
AQ


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.75
Mean = 2.9
N = 77.00

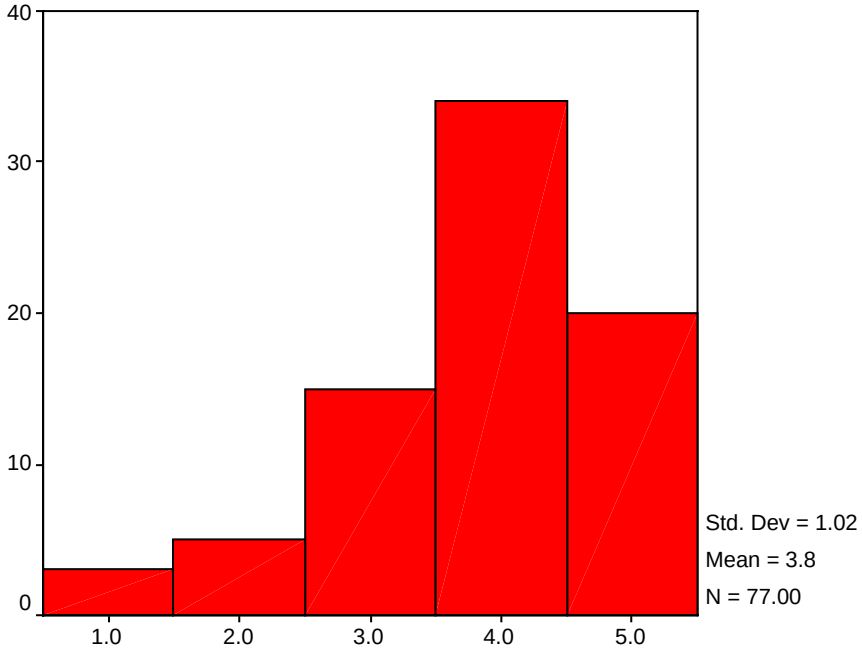
AQ


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.35
Mean = 3.4
N = 74.00

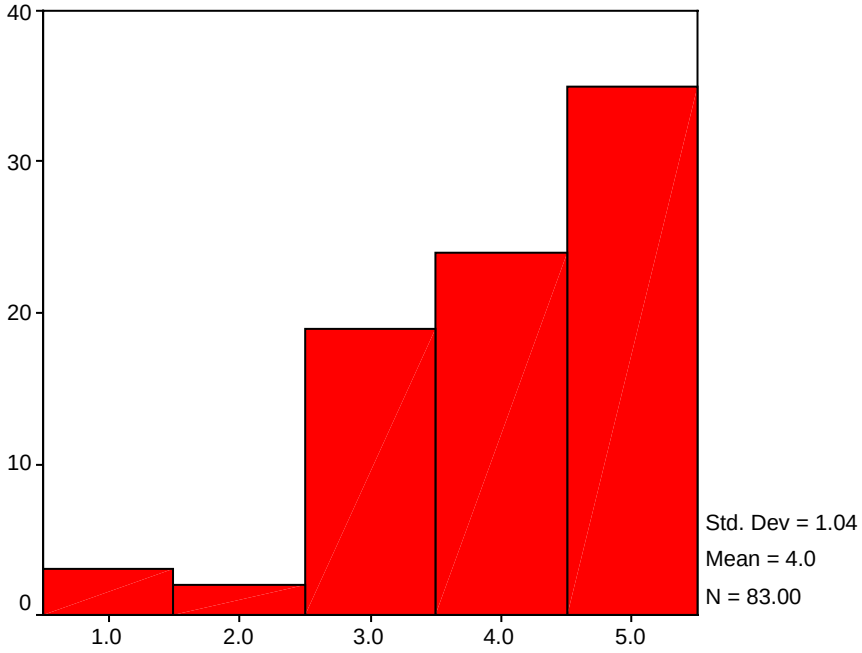
AR


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.59
Mean = 2.6
N = 82.00

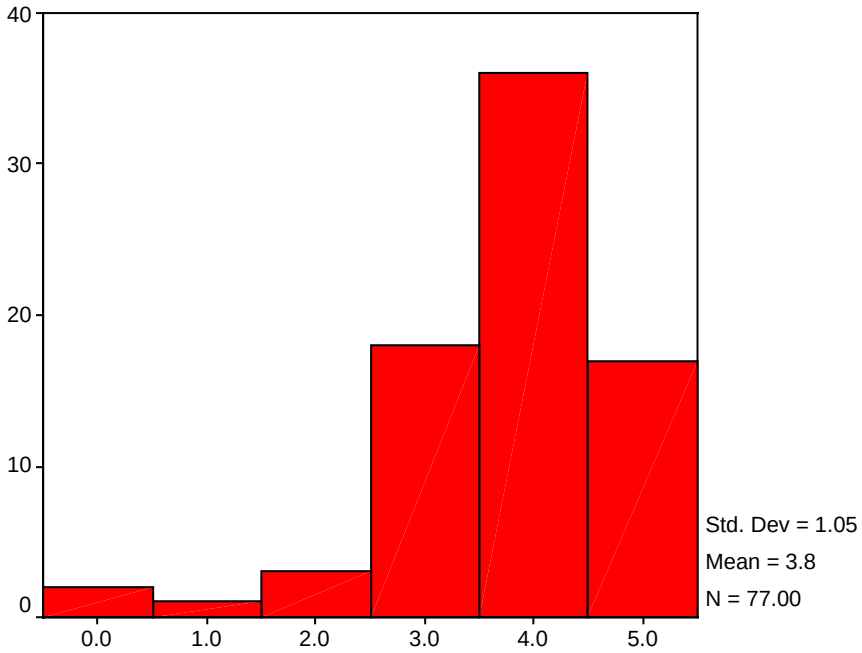
AR


40
30
20
10
0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.02
Mean = 3.8
N = 77.00

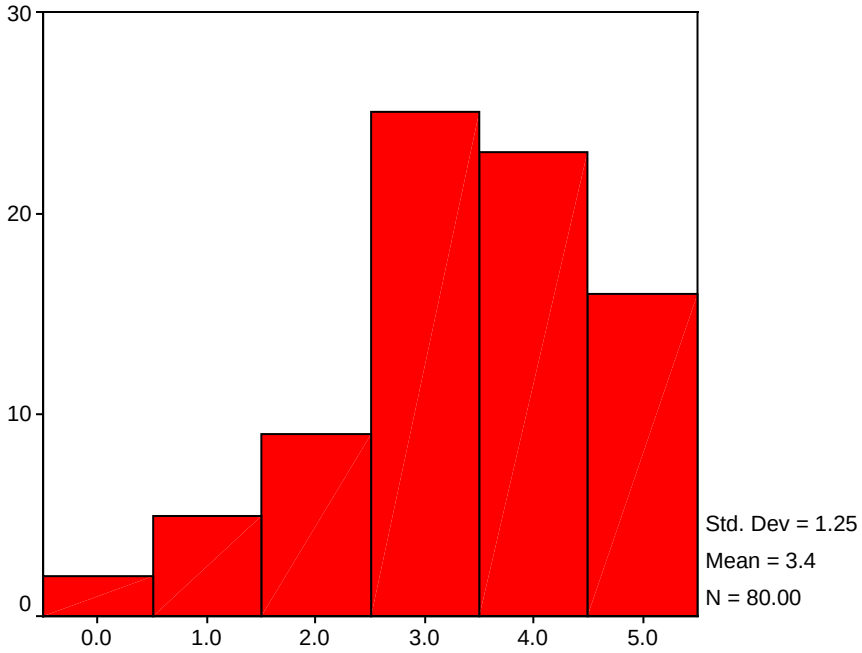
BF


40
30
20
10
0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.04
Mean = 4.0
N = 83.00

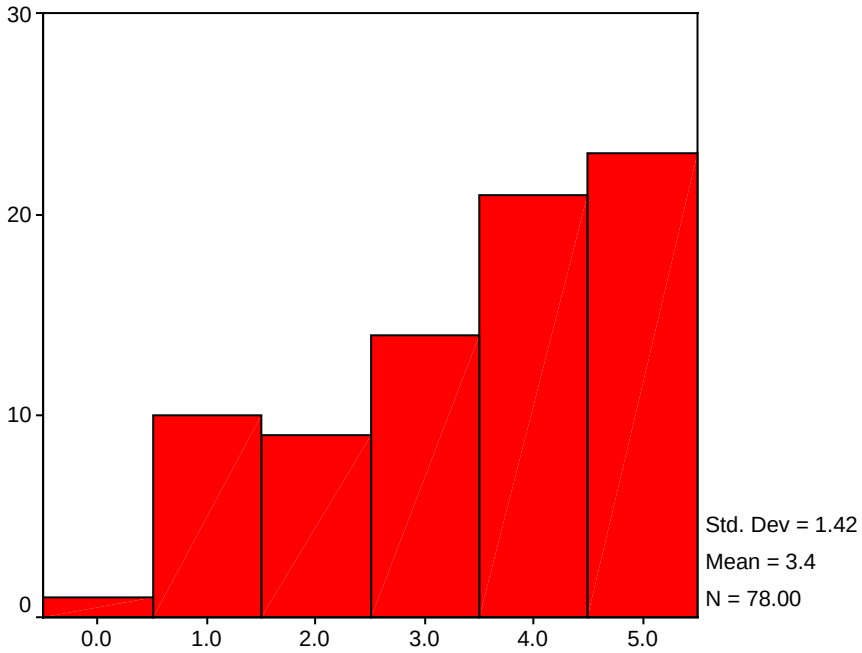
BF


40
30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.05
Mean = 3.8
N = 77.00

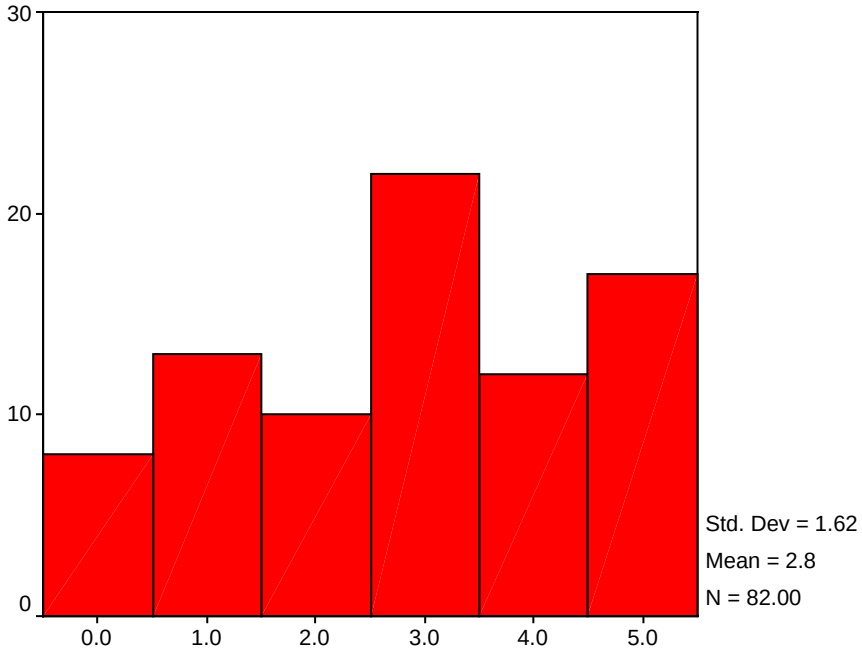
BL


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.25
Mean = 3.4
N = 80.00

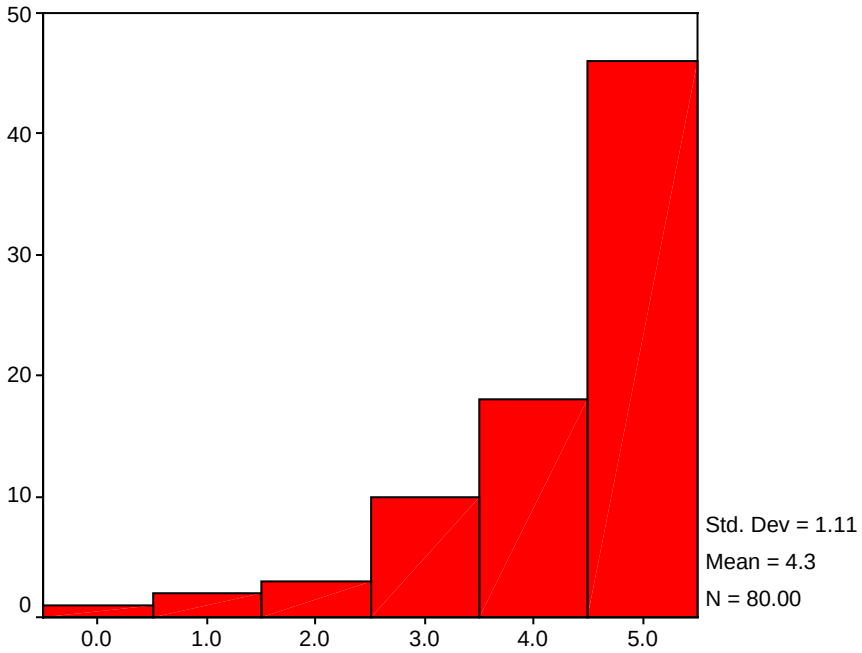
BL

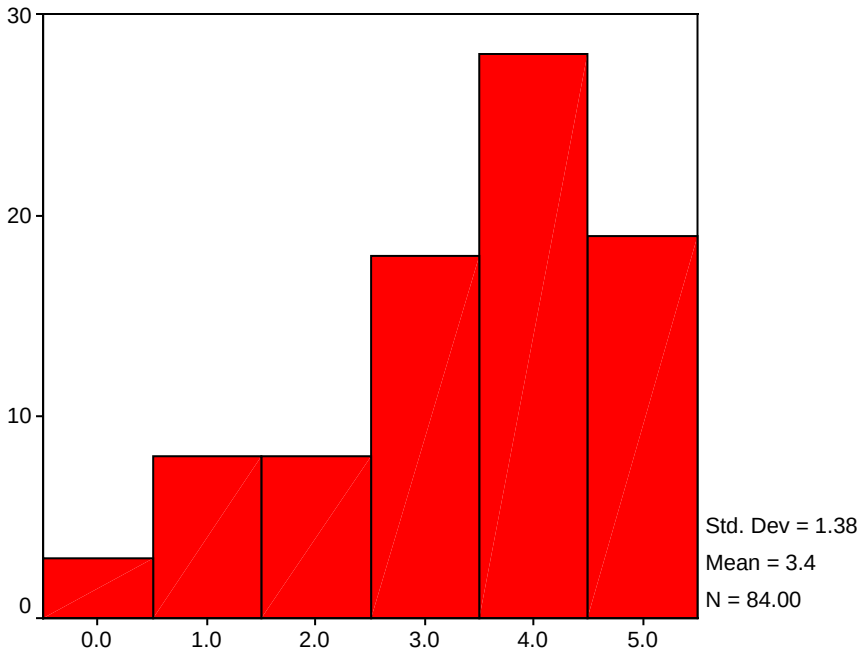

30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.42
Mean = 3.4
N = 78.00

BS


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.62
Mean = 2.8
N = 82.00

BS


50
40
30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.11
Mean = 4.3
N = 80.00
BT


30
20
10
0
0.0
1.0
2.0
3.0
4.0
5.0
Std. Dev = 1.38
Mean = 3.4
N = 84.00
BT

Appendix E

Responses of Educators and Public compared using means of Indexed Scores. Indexed Scores are calculated from original data as shown. Index range is 100 points.

Original score	5	4	3	25	1
Indexed score	100	75	50	2	0

Mean indexed scores are calculated for responses to matched question items for Educators and Public. A difference of 7.6 or greater between matched means is significant.

Question number	Educator score	Public score	Diff.	Signif.		Question number	Educator score	Public score	Diff.	Signif
16	54.6	62.2	7.6	Yes		44	89.7	90.3	0.6	No
17	70.9	65.1	5.8	No		45	79.3	89.9	10.6	Yes
18	75.3	64.1	11.2	Yes		46	60.6	59.7	0.9	No
19	41.1	31.9	9.2	Yes		47	79.8	82.7	2.9	No
20	49.0	52.1	3.1	No		48	86.9	83.7	3.2	No
21	85.8	77.7	8.1	Yes		49	76.0	69.1	6.9	No
22	76.6	74.1	2.5	No		50	44.9	55.0	10.1	Yes
23	82.1	79.4	2.7	No		51	77.6	86.3	8.7	Yes
24	82.9	71.0	11.9	Yes		52	41.7	34.7	7.0	No
25	78.9	78.8	0.1	No		53	70.8	72.6	1.8	No
26	87.8	85.7	2.1	No		54	90.6	91.1	0.5	No
27	71.8	75.9	4.1	No		55	70.5	75.9	5.4	No
28	79.8	77.1	2.7	No		56	89.9	88.3	1.6	No
29	72.5	77.1	4.6	No		57	68.2	59.6	8.6	Yes
30	73.7	71.5	2.2	No		58	64.0	64.1	0.1	No
31	60.4	67.3	6.9	No		59	81.3	86.6	5.3	No
32	88.8	92.3	3.5	No		60	70.0	68.7	1.3	No
33	73.1	79.5	6.4	No		61	71.7	61.5	10.2	Yes
34	79.0	59.4	19.6	Yes		62	85.3	89.8	4.5	No
35	80.2	71.3	8.9	Yes		63	52.5	55.3	2.8	No
36	79.1	76.8	2.3	No		64	81.1	78.4	2.7	No
37	91.5	87.2	4.3	No		65	84.2	87.1	2.9	No
38	70.6	66.2	4.4	No		66	74.0	82.0	8.0	Yes
39	79.1	68.8	10.3	Yes		67	29.2	44.3	15.1	Yes
40	61.3	62.2	0.9	No		68	62.3	53.4	8.9	Yes
41	62.0	51.4	10.6	Yes		69	82.6	63.0	19.6	Yes
42	74.4	80.1	5.7	No		70	87.3	78.7	8.6	Yes
43	84.1	84.9	0.8	No		71	75.3	75.6	0.3	No

Appendix F

Data Reduction by Factor Analysis:

Factorial analysis of data from 56 variables which compose the descriptors in the opinions of Educators and the Public as determined by the questionnaire survey instrument.

K-M-O Index (Kaiser-Meyer-Olkin) of Sampling Adequacy = 0.657

(0.5 is acceptable; 0.6 is mediocre; 0.8 is meritorious)

Bartlett's Test of Sphericity

Chi squared =3041.7 df = 1540 Significance = 0.000

The data passed these two necessary conditions and therefore were acceptable for Factor Analysis. Analysis produced 17 Factors. The highest/first ranked 10 were considered. This corresponded to Factors with Eigen values greater than 1.0 and which cumulatively could explain most of the variance. (Amount explained = 57%)

Appendix G

Naming the Factors:

The factors were named by considering those questions which yielded the highest loadings for that factor and trying to summarize them within the name. This is a very subjective exercise.

Factor 1	Learning Environment
Factor 2	Teaching Standards
Factor 3	Teaching Sensitivity
Factor 4	Societal Attitudes
Factor 5	Teacher Incentives
Factor 6	Preparation for Work
Factor 7	Publicize Performance
Factor 8	Public versus Private Education Issues
Factor 9	Teacher Attitudes
Factor 10	Student Performance Rewards

Appendix H

ANOVA analysis using Demographic data as Independent Variable (IV) versus Factor Scores as Dependent Variables (DV).

These are the Factor Scores using the first 10 ranked Factors. The Factors are identified by numbers 1 to 10. See Appendix G for the naming of the Factors.

One way ANOVA analysis of Factor Scores versus Respondent Type

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
REGR factor score 1 for analysis 1	Between Groups	3.751	1	3.751	3.856	.052
	Within Groups	98.249	101	.973		
	Total	102.000	102			
REGR factor score 2 for analysis 1	Between Groups	.756	1	.756	.754	.387
	Within Groups	101.244	101	1.002		
	Total	102.000	102			
REGR factor score 3 for analysis 1	Between Groups	2.182E-02	1	2.182E-02	.022	.883
	Within Groups	101.978	101	1.010		
	Total	102.000	102			
REGR factor score 4 for analysis 1	Between Groups	4.971	1	4.971	5.174	.025
	Within Groups	97.029	101	.961		
	Total	102.000	102			
REGR factor score 5 for analysis 1	Between Groups	16.863	1	16.863	20.005	.000
	Within Groups	85.137	101	.843		
	Total	102.000	102			
REGR factor score 6 for analysis 1	Between Groups	6.741E-02	1	6.741E-02	.067	.797
	Within Groups	101.933	101	1.009		
	Total	102.000	102			
REGR factor score 7 for analysis 1	Between Groups	.459	1	.459	.456	.501
	Within Groups	101.541	101	1.005		
	Total	102.000	102			
REGR factor score 8 for analysis 1	Between Groups	2.847	1	2.847	2.900	.092
	Within Groups	99.153	101	.982		
	Total	102.000	102			
REGR factor score 9 for analysis 1	Between Groups	.226	1	.226	.224	.637
	Within Groups	101.774	101	1.008		
	Total	102.000	102			
REGR factor score 10 for analysis 1	Between Groups	1.988	1	1.988	2.007	.160
	Within Groups	100.012	101	.990		
	Total	102.000	102			

One way ANOVA analysis of Factor Scores versus Gender

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
REGR factor score 1 for analysis 1	Between Groups	.677	1	.677	.673	.414
	Within Groups	96.522	96	1.005		
	Total	97.199	97			
REGR factor score 2 for analysis 1	Between Groups	9.839E-04	1	9.839E-04	.001	.975
	Within Groups	99.495	96	1.036		
	Total	99.496	97			
REGR factor score 3 for analysis 1	Between Groups	.951	1	.951	1.108	.295
	Within Groups	82.393	96	.858		
	Total	83.344	97			
REGR factor score 4 for analysis 1	Between Groups	1.712	1	1.712	1.724	.192
	Within Groups	95.311	96	.993		
	Total	97.023	97			
REGR factor score 5 for analysis 1	Between Groups	7.012	1	7.012	7.355	.008
	Within Groups	91.521	96	.953		
	Total	98.533	97			
REGR factor score 6 for analysis 1	Between Groups	1.258E-03	1	1.258E-03	.001	.972
	Within Groups	99.305	96	1.034		
	Total	99.306	97			
REGR factor score 7 for analysis 1	Between Groups	1.249	1	1.249	1.284	.260
	Within Groups	93.358	96	.972		
	Total	94.607	97			
REGR factor score 8 for analysis 1	Between Groups	5.075E-02	1	5.075E-02	.052	.820
	Within Groups	93.124	96	.970		
	Total	93.175	97			
REGR factor score 9 for analysis 1	Between Groups	.848	1	.848	.824	.366
	Within Groups	98.836	96	1.030		
	Total	99.684	97			
REGR factor score 10 for analysis 1	Between Groups	.373	1	.373	.388	.535
	Within Groups	92.412	96	.963		
	Total	92.786	97			

One way ANOVA analysis of Factor Scores versus Age Groups

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
REGR factor score 1 for analysis 1	Between Groups	4.094	4	1.024	1.025	.399
	Within Groups	97.906	98	.999		
	Total	102.000	102			
REGR factor score 2 for analysis 1	Between Groups	3.201	4	.800	.794	.532
	Within Groups	98.799	98	1.008		
	Total	102.000	102			
REGR factor score 3 for analysis 1	Between Groups	2.209	4	.552	.542	.705
	Within Groups	99.791	98	1.018		
	Total	102.000	102			
REGR factor score 4 for analysis 1	Between Groups	.772	4	.193	.187	.945
	Within Groups	101.228	98	1.033		
	Total	102.000	102			
REGR factor score 5 for analysis 1	Between Groups	7.080	4	1.770	1.827	.130
	Within Groups	94.920	98	.969		
	Total	102.000	102			
REGR factor score 6 for analysis 1	Between Groups	.520	4	.130	.125	.973
	Within Groups	101.480	98	1.036		
	Total	102.000	102			
REGR factor score 7 for analysis 1	Between Groups	3.624	4	.906	.902	.466
	Within Groups	98.376	98	1.004		
	Total	102.000	102			
REGR factor score 8 for analysis 1	Between Groups	8.800	4	2.200	2.313	.063
	Within Groups	93.200	98	.951		
	Total	102.000	102			
REGR factor score 9 for analysis 1	Between Groups	8.934	4	2.233	2.352	.059
	Within Groups	93.066	98	.950		
	Total	102.000	102			
REGR factor score 10 for analysis 1	Between Groups	3.192	4	.798	.792	.533
	Within Groups	98.808	98	1.008		
	Total	102.000	102			

One way ANOVA analysis of Factor Scores versus site of K-12 education

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
REGR factor score 1 for analysis 1	Between Groups	2.193	4	.548	.533	.712
	Within Groups	99.794	97	1.029		
	Total	101.986	101			
REGR factor score 2 for analysis 1	Between Groups	1.859	4	.465	.452	.771
	Within Groups	99.845	97	1.029		
	Total	101.704	101			
REGR factor score 3 for analysis 1	Between Groups	2.518	4	.629	.615	.653
	Within Groups	99.244	97	1.023		
	Total	101.762	101			
REGR factor score 4 for analysis 1	Between Groups	2.402	4	.601	.585	.674
	Within Groups	99.588	97	1.027		
	Total	101.991	101			
REGR factor score 5 for analysis 1	Between Groups	4.046	4	1.011	1.006	.408
	Within Groups	97.531	97	1.005		
	Total	101.577	101			
REGR factor score 6 for analysis 1	Between Groups	3.441	4	.860	.850	.497
	Within Groups	98.195	97	1.012		
	Total	101.635	101			
REGR factor score 7 for analysis 1	Between Groups	2.780	4	.695	.680	.608
	Within Groups	99.181	97	1.022		
	Total	101.961	101			
REGR factor score 8 for analysis 1	Between Groups	4.279	4	1.070	1.062	.380
	Within Groups	97.702	97	1.007		
	Total	101.981	101			
REGR factor score 9 for analysis 1	Between Groups	8.494	4	2.124	2.204	.074
	Within Groups	93.455	97	.963		
	Total	101.949	101			
REGR factor score 10 for analysis 1	Between Groups	10.750	4	2.688	2.864	.027
	Within Groups	91.016	97	.938		
	Total	101.766	101			

One way ANOVA analysis 0f Factor Scores versus Highest educational level

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
REGR factor score 1 for analysis 1	Between Groups	1.227	4	.307	.278	.892
	Within Groups	95.065	86	1.105		
	Total	96.293	90			
REGR factor score 2 for analysis 1	Between Groups	3.684	4	.921	.856	.494
	Within Groups	92.526	86	1.076		
	Total	96.210	90			
REGR factor score 3 for analysis 1	Between Groups	.890	4	.222	.239	.915
	Within Groups	79.965	86	.930		
	Total	80.855	90			
REGR factor score 4 for analysis 1	Between Groups	4.730	4	1.182	1.118	.353
	Within Groups	90.947	86	1.058		
	Total	95.677	90			
REGR factor score 5 for analysis 1	Between Groups	2.644	4	.661	.642	.634
	Within Groups	88.512	86	1.029		
	Total	91.156	90			
REGR factor score 6 for analysis 1	Between Groups	10.889	4	2.722	2.913	.026
	Within Groups	80.360	86	.934		
	Total	91.249	90			
REGR factor score 7 for analysis 1	Between Groups	4.270	4	1.067	1.118	.354
	Within Groups	82.144	86	.955		
	Total	86.414	90			
REGR factor score 8 for analysis 1	Between Groups	8.894	4	2.223	2.261	.069
	Within Groups	84.578	86	.983		
	Total	93.471	90			
REGR factor score 9 for analysis 1	Between Groups	4.321	4	1.080	1.061	.381
	Within Groups	87.571	86	1.018		
	Total	91.892	90			
REGR factor score 10 for analysis 1	Between Groups	3.935	4	.984	1.002	.411
	Within Groups	84.431	86	.982		
	Total	88.366	90			

One way ANOVA analysis of Factor Scores versus Time of last educational event

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
REGR factor score 1 for analysis 1	Between Groups	4.065	3	1.355	1.303	.278
	Within Groups	94.619	91	1.040		
	Total	98.684	94			
REGR factor score 2 for analysis 1	Between Groups	2.539	3	.846	.812	.491
	Within Groups	94.896	91	1.043		
	Total	97.435	94			
REGR factor score 3 for analysis 1	Between Groups	6.947	3	2.316	2.489	.065
	Within Groups	84.667	91	.930		
	Total	91.614	94			
REGR factor score 4 for analysis 1	Between Groups	2.343	3	.781	.738	.532
	Within Groups	96.279	91	1.058		
	Total	98.621	94			
REGR factor score 5 for analysis 1	Between Groups	5.584	3	1.861	1.969	.124
	Within Groups	86.041	91	.946		
	Total	91.625	94			
REGR factor score 6 for analysis 1	Between Groups	6.697	3	2.232	2.336	.079
	Within Groups	86.970	91	.956		
	Total	93.667	94			
REGR factor score 7 for analysis 1	Between Groups	2.705	3	.902	.914	.438
	Within Groups	89.765	91	.986		
	Total	92.470	94			
REGR factor score 8 for analysis 1	Between Groups	1.578	3	.526	.518	.671
	Within Groups	92.486	91	1.016		
	Total	94.065	94			
REGR factor score 9 for analysis 1	Between Groups	1.522	3	.507	.515	.673
	Within Groups	89.680	91	.985		
	Total	91.202	94			
REGR factor score 10 for analysis 1	Between Groups	5.405	3	1.802	1.764	.160
	Within Groups	92.919	91	1.021		
	Total	98.324	94			

Bibliography

Alexander, F. King (2000). The changing face of accountability. The Journal of Higher Education. Vol.7 No.4. p.411 – 431.

Allen, Jeanne (1994). A primer on school accountability. Center for Education Reform. Retrieved May 12, 2001 from http://edreform.com/pubs/account1.htm

Alfred, R.; Ewell, P.; Hudges, J. and McClenney, K. (1999). Core indicators of effectiveness for community colleges. Washington, D. C. Community College Press.

Alreck, P. L. and Settle, R. B. (1995). The survey research handbook. 2nd. Edition. Guidelines and strategies for conducting a survey. Burr Ridge, Ill. Irwin Professional Publishing.

Alvarez, L. (2001). Senate passes Bill for immediate tests in public schools. New York Times June 15, 2001 pA1

.

American Association of Community Colleges. (1999). Final Draft. A report of the Community College Roundtable. Community Colleges: Core Indicators of Effectiveness. Community College Press. Washington, D. C.

American Council on Education (2001). Remarks before the American Council on Education. U. S. Secretary of Education Rod Paige. Feb. 20, 2001. Accessed July 13, 2001from http://www.ed.gov/speeches/02-2001/010220.html

American Council on Education (1998). Straight talk about college costs and prices. (Report of the National Commission on the Cost of Higher Education)

American Productivity and Quality Center (2000). Benchmarking Best Practices in Accountability Systems: Executive Summary.

American Productivity and Quality Center (2001). The customer-centric contact center: A new model. Retrieved May 2001 from http://www.apqc.org/free/articles/disp

American Productivity and Quality Center International Benchmarking Clearinghouse (2001). (i) Baldrige-based self-assessments (ii). Measure what matters: Aligning performance measures with business strategy (iii) Quality approaches for the new Millenium. (iv) Achieving organizational excellence through the performance measurement system (v) Process clarification framework (vi) Establishing balanced scorecards (vii)Six Sigma going beyond manufacturing (8)A new approach to assessing benchmarking progress. Retrieved May 2, 2001 from http://www.apqc.org/free/articles/disp

American Productivity and Quality Center (2001). (i).Measuring Institutional Performance Outcomes (ii)Corporate Performance Measures (iii) Focus on K-16 education: Improvement Tools. Retrieved on Feb. 9, 2001 from http://www.apqc.org

Association of Colleges of Applied Arts and Technology of Ontario (1996). Accountability in a learning-centred environment: Using Performance Indicators in Reporting, Advocacy and Planning.

Atkinson-Grosjean, J. and Grosjean, G. (2000). The use of performance models in higher education: A comparative International review.

Babbie, Earl (1999). The Basics of Social Research. Belmont, Ca. Wadsworth Publishers.

Baldrige National Quality Program (2001). (i) Criteria for performance excellence (ii) Encouraging performance excellence (iii) Why apply? (iv) Education criteria for performance excellence.

Barnetson, Robert J. (1997). Changing definitions and off-loading responsibility in Alberta's post-secondary system. Education Policy Analysis Archives Vol5 Number 22.

Barton, Paul E. (2001). The Elementary and Secondary Education Act and Standards-Based Reform. Center on Education Policy. Feb. 2001. Retrieved June 11, 2001 from http://www.ctredpol.org

Beering, S. C. (2000).Purdue University: Excellence 21. Accessed via Excellence Homepage.

Behnhe, Patricia C. (2001). Accountability in the public schools? Accesses Aug. 10, '01 from http://www.5ire.com/article.cfm/public_schools-fam

Bello, R. (2000). Officals address lagging Hispanic education rates. The Daily Texan. Sept. 27, 2000. p.1, 2.

Bennett, W. E. and Finn, C. E., Jr. (2000). The real improvement in Texas schools. The New York Times. (Oct. 20, 2000).

Bermuda Government (1996). A curriculum management audit of the Bermuda Government Public School Sytsem conducted by the International Curriculum Management Audit Centre.

Bermuda Government (1999). Bermuda Digest of Statistics 1999. Published by Ministry of Finance.

Bermuda Government (2002). Ministry of Education and Development. Budget brief 2002-2003.

Bermuda Ministry of Education (1989). Final report of the Education Planning Team.

Bermuda Ministry of Education (1994). (i) Restructuring: a progress report. (ii) Restructuring: questions and answers.

Bermuda Ministry of Education (2000). External analysis and audit of selected organization and system attributes of the Bermuda College.

Berry, L. L. ; Parasuraman, A. and Zeithaml, V. A. (1994). Improving service quality in America: Lessons learned. Academy of Management Executive Vol. 8 No.

Black, Earl and Black, Merle (1987). Politics and society in the South. London: Howard University Press.

Bowers, Bruce C. (1989). State-enforced accountability of local school districts ERIC Digests ED309556. Retrieved June 10, 2001 from http://www.edgov/databases/ERIC Digests/ed309556.html

Bradshaw, Michael D. (2001). A structural comparison of some statewide accountability system reports for K-12. (Unpublished).

British Broadcasting Corporation News (1999). Education: School league tables . Explanation of the figures. Primary schools England 1998. Retrieved Feb. 9, 2001 from http://news.bbc.co.uk/hi/english/static/education/school_tables_1998/key_stag…/explanation.htm

Bush, George W. (1998). Making Texas a beacon state.

Business Roundtable, The (1996). A business leader's guide to setting academic standards. Accessed June 16, 2001 from http://www.brtable.org/documents.cfm/461

Business Roundtable, The (1998). Education (1).Building support for tests that count.

Business Roundtable ,The (2000). What parents, students and teachers think about standards, tests, accountability … and more.

Cabinet Office (2000). The Six Standards for Central Government. (Effective Performance Division). Retrieved Feb. 9, 2001 from http://www.servicefirst.gov.uk/2000/introduc/six.htm

California Assessment Institute, The (2001). Resources for the Assessment Institute. Retrieved on May 17, 2001 from http://www.ca~assessment~inst.org/Resources/Resources.htm

California Dept. of Education (2000). Reporting the academic performance index growth and awards for 1999-2000. Public Schools Accountability Act. From http://www.cde.ca.gov/psaa

Callan, Patrick M., Doyle, William and Finney, Joni E. (2001). Measuring up 2000. Evaluating state higher education performance. Change March/April 2001.

Center for Education Reform Acton Paper (2000). The new generation of standardized testing. Retrieved Dec. 2, 2000 from http://edreform.com/pubs/testing.htm

Chronicle of Higher Education, The (March 2, 2001). Scholars say high-stakes test deserve a failing grade. D. W. Miller p.A14.

Clark, C. (1998). What we know about student performance in Texas. Texas Center for Educational Research. Retrieved Sept. 30, 2000 from Texas Association of School Boards at http://tasb.org/schools/performance.html

Clark,C. T. and Schkade, L. L. (1974). Statistical analysis for administrative decisions. Cincinnati. South-Western Publishing Co.

Cohen, A. M. and Brawer, F. B. (1996). The American community college. 4th. Ediction. San Francisco. Jossey Bass Inc.

Council of Chief State School Officers (1999). State education accountability reports and indicator reports. Status of reports across the states 1999.

Council of Chief State School Officers (2000). State education accountability systems.

Council of Great City Schools (2000). Proposal to conduct trial NAEP assessment for large urban school districts.

Creswell, John W. (1998). Qualitative inquiry and research design - - choosing among 5 options. Thousand Oaks, California. Sage Publishers Inc.

Cunningham, L. L. (1972). Our accountability problems. In 'Accountability in American education.' Edited Sciara, F. J. and Jantz, R. K. Boston, Mass. Allyn and Bacon Publishers, Inc.

Cuttance, P. and Stokes, S. A. (2000). Reporting on student and school achievement. Research report for Commonwealth Department of Education, Training and Youth Affairs, Commonwealth of Australia.

Daniel, W. W. (1983). Biostatistics: A foundation for anaysis in the health sciences. 3rd. Edition. N.Y. John Wiley and sons.

Darling-Hammond, L and Ascher, C. (1991). Accountability mechanisms in big city school systems ERIC Digests ED 334311. Retrieved Oct. 12, 2000 from http://www.ed.gov/databases/ERIC_Digests/ed334311.html

Davey, Lynn (!992). The case for a national testing system. ERIC Digests ED410239. Retrieved June 10. 2001 from http://edgov/databases/ERIC Digests/ed410239.html

Deming, (1993). Out of the crisis. Cambridge, Mass. Center for Advanced Engineering Studies.

Dennis, Dion (1995). Brave new reductionism: TQM as ethnocentrism. Education

Department of Further Education in England (2001). Education and training 16-25. Retrieved Jan. 26, 2001 from http://www.dfee.gov.uk//etb/chapter3.htm

Dill, David D. (1999). Are academic audits the answer? "Assessment movement" seeks to measure and improve student learning. National Cross Talk Vol. 7 No. 3. Summer 1999.

Dillman, D. A. (1978). Mail and telephone surveys; the total design method. NY, NY. Wiley.

Dorn, Sherman (1998). The political legacy of school accountability systems. Education Policy Archives Vol. 6 No.1 .

Doyle, Denis P. (1996). The Education Summit: What does it mean for education reform? Thomas Fordham Foundation. Retrieved October 18, 2000 from http://edexcellence.net/issuespl/subject/standar/doyle.html

Drucker, Peter F. (1999). Management Challenges for the 21st. Century. HarperCollins Publishers, Inc. N.Y., N.Y.

Edwards, T. L. (2000). Impact of future trends and issues on the role of departmental heads in Canadian community colleges. Doctoral thesis. University of Texas at Austin.

ERIC ED429988. (1999). A nation still at risk. ED 429988.

Ewell, Peter T (2001). Statewide testing in higher education. Change. March/April 2001.

Ewell, Peter T. (1998). Examining a brave new world: How accreditation might be different. Retrieved from Council for Higher Education Accreditation on Feb. 15, 2001 at http://www.chea.org/Events/Usefulness/98May/98_05Ewell.html

Finn, C. E. Jr. (1999). Getting serious about the schools. Retrieved from Thomas B. Fordham Foundation on Oct. 18, 2000 at http://www.edexcellence.net/library/gopedu.html

Finn, C. Jr. and Petrillili, M. (Ed.) 2000. The state of state standards . Thomas B. Fordham Foundation. Washington, D.C.

Florida Dept. of Education (2000). (i) Educator accountability. (ii)School accountability. Retrieved from http://firn.edu/doe/bin00048/basics/school1.htm

Fowler, F. F. (2000). Policy studies for educational leaders: An introduction. Upper Saddle River, N. J. Prentice-Hall Inc.

Fuchs, L. H. (1990). The American kaleidoscope: Race, ethnicity and the civic culture.

Fuhrman, S. H. (1999). The new accountability. Consortium for Policy Research in Education. RB-27-January, 1999.

Fuller, Edward J and Johnson, Jr., Joseph F. (2001). Can state accountability systems drive improvements in school performance for children of color and children from low-income homes? Education and Urban Society. Vol 33 (3). May 2001. p. 260-283.

Garvin, David A. (1983). Competing on the eight dimensions of quality. Pre publication article from the book "Managing Quality" by Free Press.

Gifford, M. (1998). Charter schools. Double the accountability. Accessed April 14, 2001 from wysiwyg://134/http://www.goldwaretinstitute.org/perspectives/1998/9815.htm

Gifford, Mary (2000). How to really accomplish accountability in public education (without a sales tax increase) Retrieved from wysiwyg://120/http://www.goldwaterinstitute.org/perspectives/2000/0022.htm

Goff, John M. (2000). A more comprehensive accountability model. Council for Basic Education. Retrieved January 18, 2001 from http://www.c-b-e.org/articles/goff.htm

Goodenough, A. (2000) New program will provide NY public schools with a head start on college. New York Times. June 7, 2000. pC19.

Greenhouse, D. (2000). Chance for change, if city and teachers agree. New York Times June 5, 2001. pA24.

Hacker. A (1995). Two nations. N.Y. Ballantine Books.

Haney, W. (2000). The myth of the Texas miracle in education. Education Policy Analysis Archives Vol.8 number 41.

Hart, C. W. (1988). The power of unconditional service guarantees. HBR. July/Aug. 1988.

Hauser, J. R. and Clausing, D. (1988). The house of quality. HBR. May-June 1988.

Hawley, Willis D. and Valli, Linda (1998). Guide to the National Partnership for excellence and accountability in Teaching (NPEAT). ERIC Digests ED426056. Retrieved June 10, 2001 from http://www.edgov/databases/ERIC Digests/ed426056.html

Heinze, D. (1980). Fundamentals of managerial statistics. Cincinnati. South-Western Publishing Co.

Horsch, K. (1996). Results based accountability systems: opportunities and challenges. Harvard Family Research Project Vol. 2. Number 1.

Hussey, Jill and Hussey, Roger (1997). Business research. London. MacMillan Press.

Immerwahr, J. (2000). Doing comparatively well: Why the public loves higher education and criticizes K-12. National Cross Talk Vol 8. No. 1. of the National Center for Public Policy and Higher Education. Retrieved on October 7, 2000 from http://www.highereducation.org/crosstalk/ct0100/voices0100~immerwahr.shtml

Internal Audit Branch Office of the ComptrollerGeneral Ministry of Finance and Corporate Relations (1996). Accountability in the education systems. Retrieved from the British Columbia Ministry of Education on Jan. 25, 2000 at http://www.bced.gov.bc.ca/accountability

Jacobs, J. (2001). Texas students make gains in Maths. Austin American Statesman. Aug. 3, 2001. pA1.

Jennings, Jack (2001). An education agenda for the Congress and the New Administration. Retrieved June 10, 2001 from http://www.ctredpol.org/

Jerald, C. D. (2001). Real results, remaining challenges: The story of Texas education reform. Published by The Business Roundtable.

Johns, Gary (1996). Organizational Behavior; Understanding and managing life at work. 4th. Edition. HarperCollins College Publishers, Inc. N.Y., N.Y.

Johnson, J. F. Jr., (1999). Improving public schools in Texas. Basic Education October 1999. Retrieved Jan. 18, 2001 from http://www.c-b-e.org/articles/texas.htm

Juran, Joseph M. (1993). Made in U.S.A.: A renaissance in quality. Harvard Business Review (July-August) 1993.

King, Sharon H. (2001). High school/college collaborative: The impact of a co-enrollment program on student success. Doctoral thesis.

Klein, S.; Hamilton, L.; McCaffrey, D and Stecher, B. (2001). What do test scores in Texas tell us? Retrieved Jan. 18, 2001 from Rand Education at http://www.rand.org/publications/IP/IP202/

Kozol, J. (1991). Savage inequalities: Children in America's schools. NY, NY. Crown Publishers.

Kruger, W. S. (1972). Implications of accountability for educational program evaluation. In 'Accountability in education in American education.' Edited by Sciara, F. J. and Kantz, R. K. Boston, Mass. Allyn and Bacon Publishers.

Ladd, Helen F. (1996). (Ed.) Holding school accountable. Performance-based reform in education. Washington, D. C. Published by Brookings Institute.

Lahiri, Jaideep (1999). The enigma of Six Sigma. India Today (Business Today) Sept. 22nd. 1999. Retrieved on December 2, 2000 from vysiwyg://245/http://www.India-today.com/btoday/19990922/cover.html

Lashway, Larry (1999). Holding schools accountable for achievement. ED 434381. Educational Research and Improvement Clearinghouse . Retrieved June 10, 2001 from http://www.edgov/databases/ERIC-Digests/ed434381.html

Lessinger, L. M. (1970). Accountability in education: Every child a winner. NY, NY. Simon and Schuster.

Lynd, C. (2000) (1990??) The new generation of standardized testing. Center for Education Reform. Retrieved Dec. 2000 from http://edreform.com/pubs/testing.htm

Manno, B.V. (1998). How to fix a nation still at risk. Retrieved from Thomas B. Fordham Foundation on Oct. 18, 2000 at http://edexcellence.net/library/fixnsar.html

Matter, Kevin M. (1999). Strategies for improving the process of educational assessment. ERIC Clearinghouse on Assessment and Evaluation. EDO-TM-99-03.

McFadden, F. R. (1993). Six Sigma Quality Programs. Quality Progress. June, 1993.

Meister, Jeanne C. (1994). Corporate quality universities. N. Y., N. Y. Published by Irwin Professional Publishing. .

Menzes, M. A. and Serbin, J. (1993). Xerox Corporation: The Customer's satisfaction program. Boston, Mass. Published by Harvard Business School.

Missouri Dept. of Higher Education (1999). Blueprint for Missouri Higher Education. Reporting the results. Annual Report Coordinating Board for Higher education.

Mitchell, K. J. (1996). Reforming and conforming NASDC principals discuss school accountability systems. Retrieved 2000 from http://www.rand.org/publications/MR/MR716/

Mize, R. M. (1999). Accountability measures: A comparison by type and state. Sacramento, CA. Published by Community College League of California.

Moore, W. Jr. (1971). Blind man on a freeway. San Francisco. Jossey-Bass, Inc.

Motorola's Six sigma Program: An overview. Retrieved on December 2, 2000 from http://web.wt.net/~jquick/sixsigma.html

National Association of College and University Business Officers (1999). The 1998-1999 NACUBO Benchmark Program. Retrieved Oct. 26, 2000 from http:www.nacubo.org/website/benchmarkingsmall.html

National Center for Education Statistics (2000). International Education Indicators: A time series perspective, 1985-1995. Washington, D. C. Published by U. S. Department of Education.

National Cross Talk (1999). "Academic Audits". American educators examine European practices. National Cross Talk Vol.7 No.3 Summer 1999. The National enter for Public Policy and Higher Education.

National Education Goals Panel (1997). Implementing academic standards.

National Education Goals Panel (1998). Talking about tests: an idea book for state leaders. Washington, D. C.: U. S. Government Printing Office.

National Education Goals Panel (1999). 10th Anniversary: Building on the momentum.

National Education Goals Panel (1999). Data Volume for the National Education Goals Report.

National Education Goals Panel (1999). Promising practices: progress towards the Goals.

National Education Goals Panel (2000). Minnesota and TIMSS: Exploring high achievement in eight grade Science.

National Education Goals Report (1998). Building a nation of learners.

National Education Goals Report Summary (1997). Mathematics and Science Achievement for the 21st. Century.

National Partnership for Reinventing Government (1999). Balancing measures: Best practices in performance management.

Northcutt, N. (2001). Class notes.

OECD/CERI (2000). Schooling for tomorrow. ICT and the quality of learning. Methodology for case studies of organizational change. Retrieved June, 30, 2001 from http://bert.eds.udel.edu/oecd/cases/CASES11.html

Olson, Lynn (2001). Overboard on testing. Editorial Projects in Education Vol 20 Number 17. Retrieved Jan. 18, 2001 from http://www.edweek.org/sreports/qc01story.cfin?slug=17test.h20

Paige, Rod (2001). Remarks before the American Council on Education. Retrieved from http://www.ed.gov./Speeches/02-2001/010220.html

Parker, Mike and Slaughter, Jane (1994). Beware! TQM is coming to your campus. National Education Association. Higher Education: Thought and action TQM. Spring 1994.

Patton, Michael Q. (1990). Qualitative education and research methods. Second edition. Newbury Park, California. Sage Publishers.

Pedhazur, E. J. and Schmelkin, L. P. (1991). Measurement, design and analysis. Hillsdale, N. J. Lawrence Erlbaum Associates.

Polley, J. W. and others (1972). Problems associated with equalization of educational opportunity. In 'Accountability in American education'. Edited by Sciara, F. and Jantz, R. Boston, Mass. Allyn and Bacon.

Public Schools Accountability Act of California. (2000). Retrieved from http://www.cde.ca.gov/psaa

Rahn, M. L. and Holmes, P. (1999). Accountability systems: performance standards and assessment. NCRVE #4. March 1999.

RAND Research Brief (1999). Test-based accountability systems: lessons of Kentucky's experiment.

Reis, R. (2002). Management fads in higher education review. From on-line journal Tomorrow's Professor @ lists.Stanford.EDU.

Rose, Lowell C. and Gallup, Alec M. (1998). The 30th. Annual Phi Delta Kappa/Gallup Poll of the Public's Attitudes towards the Public Schools. Phi Delta Kappan V.80 N1. p.41 –56.

Rose, Lowell C. and Gallup, Alec M. (1999). The 31st. Annual Phi Delta Kappa/Gallup Poll of the Public's Attitudes towards the Public Schools. Phi Delta Kappan V.81 N1 p.41 – 56.

Rose, Lowell C. and Gallup, Alec M. (2000). The 32nd. Annual Phi Delta Kappa/Gallup Poll of the Public's Attitudes towards the Public Schools. Phi Delta Kappan V.82 N1 p.41 – 58.

Rose, Lowell C. and Gallup, Alec M. (2001). The 33rd. Annual Phi Delta Kappa/Gallup Poll of the Public's Attitudes towards the Public Schools. Phi Delta Kappan V.83 N1 p.41 – 58.

Rothstein, R. (2000). Add social changes to the factors affecting declining test scores. New York Times Oct. 25, 2000. p. A25

Rothstein, R. (2001). Test only here and there and never all of the above. Daily Texan. June 6, 2001. pA28.

Roueche, J. R. (1972). Accountability in the 2-year college. In Sciara, F. J. and Jantz, R. K. Accountability in American education. Boston, Mass. Allyn and Bacon Publishers.

Roueche, J. R.; Baker, G. B. III; OmahaBoy, N. H. and Mullins, P.L. (1987). Access and excellence – the open door college. Washington, D. C. Community College Press.

Roueche, J. E.; Johnson, L. F.; Roueche, S. D. and associates (1997). Embracing the tiger: The effectiveness debate and the community college. Washington, D. C. Community College Press.

Roueche, J. E.; Ely, E. E. and Roueche, S. D, (2001). In pursuit of excellence: The Community College of Denver. Washington, D.C. Community College Press.

Ruppert, Sandra S. (1998). Focus on the customer: A new approach to state-level accountability reporting and processes for higher education. Denver, CO. Published by State Higher Education Executive Officers

Rylander, C. K. (2000). The impact of the state higher education system on the Texas economy. Special Report of the Texas Comptroller of Public Accounts.

Sandham, J. L. (2001). Quality counts 2001: Parent power. Education Week on the web. Accessed http://www.edweek.org/sreports/qc01/articles/QC01/story.cfm?slug=17benito,h20

Sanger, D. E. (2001). Bush pushes ambitious education plan. New York Times. Jan. 24, 2001. pA1.

Saul, D. J. (1973). A survey of the Bermuda educational system. Privately printed.

Schemo, D. J. (2001). Bush seems to ease his stance on the accountability of schools. New York Times. July 10, 2001. p.1.

Scheurich, J. J., Skrla, L and Johnson, J. F. (2000). Thinking carefully about equity and accountability. Phi Delta Kappan. 82(4), 293-299

Schwartz, J. D.; Geltner, P. B. and Young, D. F. (1999). Indicators of institutional effectiveness: design and implementation. Santa Monica College.

Sclafani, Susan (2001). Using an aligned system to make real progress for Texas students. Education and Urban Society. Vol 33(3). May 2001. p. 305-312.

Sims, Sebrenia J. (1992). Student outcomes assessment: A historical review and guide to program development. N.Y. and London. Greenwood Press.

Sirotnik, K. A. and Kimball, K. (1999). Standards for standards-based accountability systems. Retrieved January 25, 2001 from http://www.pdkintl.org/kappan/ksir9911.htm

Skrla, Linda and Scheurich, James J. (2001). Displacing deficit thinking in school district leadership. Education and Urban Society. Vol 33(3). May 2001. p. 235-259.

Skrla, Linda., Scheurich, James J. and Johnson, Jr. Joseph.F. (2000). Equity-driven achievement-focused school districts: A report on systemic school success in four Texas school district serving diverse student populations. Charles A. Dana Center, Univ. of Texas at Austin.

Slife, B. D. and Williams, R. N. (1995). Discovering hidden assumptions in the behavioural sciences. Thousand Oaks. Sage Publications, Inc.

Spanbauer, S. J. (1994). Reenergizing community college education through Total Quality leadership. Catalyst. Volume XXIII, number 3.

Standard and Poor's School Evaluation Services (2001). Retrieved on June 16, 2001 from http://www.ses.standardandpoors.com/Homepage.html

Stasoft (2003). Principal Components and Factor Analysis. Retrieved on Feb. 26, 2003 from http://www.stasoft.com/textbook/stafacan.html

State Board of education (1995). Long range plan for public education 1996—2000. Published by Texas Education Agency.

Strong, Karen (1999). Texas outscores all other states again! National study shows Texas is best at boosting teacher quality. Retrieved Sept. 30, 2000 from Texas Association of School Boards at http://www.tasb/org/tasb_store/lonestar/1999/annual/strong.html

Tellis, Winston (1997). Introduction to case study. The Qualitative Report Vol. 3 #2.

Texas Classroom Teachers Association (2000). TAAS passes test of time: bar being raised again. Retrieved Sept. 9, 2000 from http://www.tcta.org/legal/taasbar.htm

Texas Education Agency (1998). A report to the 76th. Texas legislature. 1998 Comprehensive Biennial Report on Texas Public Schools.

Texas education Agency (1999). A report to the 76th. Texas legislature. 1999 Interim Report on Texas Public Schools.

Texas Education Agency (1999). Snapshot '99. 1998-99 School District Profiles.

Texas Education Agency (2000). 2000 Accountability Manual. The 2000 Accountability Rating System for Texas Public Schools and School Districts.

Texas Education Agency (2000). District accountability ratings for Texas public schools districts 1999 through 2000. Retrieved Sept, 2000 from http://www.tea.state.tx.us/perfreport/account/ratngsxyrs.html

Texas Education Agency State Board of Education (1995). Long range plan for public education 1996-2000.

Texas Higher Education Coordinating Board (2000). Texas Higher Education Plan: Closing the gaps by 2015.

Texas Higher Education Coordinating Board (2001) Texas public universities. Summary of Spring 2000 Data Base.

Texas Higher Education Co-ordinating Board (2001). Agency Strategic Plan for the fiscal years 2001-2005.

Texas Higher Education Coordinating Board (2001). Texas public community and technical colleges 2000 statewide Factbook.

Texas Higher Education Coordinating Board (2001). Texas public community, technical and state colleges. Summary of Spring 2000 Data Base.

Thor, Linda M. (1994). Introducing the human side of total quality management into educational institutions. Community College Journal of Research and Practice vol 18. pp359-368.

Trombey, W. T. Jr. (2000). How does Measuring up 2000 measure up. National Crosstalk Winter 2001.

Tyack, D. (1999). Choice options. School choice, Yes – But what kind? The Maerican Prospect Vol. 10 Issue 42. Accessed Aug. 28, 2001 from http://www.prospect.org/print-friendly/print/v10/42/tyack-d.html

Tyack, D. (2001). Prof. David Tyack on democracy and public schools. Accessed Aug. 20, 2001 from wysiwyg://18/http://www.sba.org/q&a/tyack.htm

U. S. Department of Education (1992). Education Research Report: Parental satisfaction with schools and the need for standards. Retrieved from http://www.ed.gov/pubs/OR/ReseachRpts/parents.html

U. S. Department of Education (1992). Educational Research Report: Meeting goal 3: How well are we doing? Retrieved from http://www.ed.gov/pubs/ResearchRpts/goal3.html

U. S. Department of Education (1994). Education Research Report: Who's in charge? Teachers views on control over school policy and classroom practices. Retrieved from http://www.ed.gov/pubs/OR/ReseachRpts/whos.html

U.S. Department of Education (1999). Demonstrating results: An introduction to the Government Performances and Results Act. (Higher Education Programs, Office of Postsecondary Education).

University of Texas System, The (2000). Service to Texas in the new Century: A long range plan for the University of Texas System.

Valencia, Richard R.; Valenzuela , Angela; Sloan, Kris; and Foley, Douglas E. (2001). Let's treat the cause, not the symptoms: Equity and accountability in Texas revisited. Phi Delta Kappan V. 83 N.4 accessed May 13, 2002 from http://www.fairtest.org/nettest/valencia%20Exerpts.html

van Raemdonck, Dirk C. (2000). European observations on U. S. public education. Goldwater Institute. Retrieved April 14, 2001 from wysiwyg://117/http:www.goldwaterinstitute.org/perspectives/2000/0028.htm

Vaughn, G. B. and associates (1992). Dilemmas of leadership: Decision making and ethics in the community college. San Francisco. Jossey-Bass Publishers.

Viadero, D. (2000). Candidates spar over test gains in Texas. Education Week on the Web Nov.1, 2000. Retrieved Nov. 2000 from http://ed.week.org/ew/ewstory.cfm?slug=09rand.h20

Vinovskis, Maris A. (1998). Overseeing the Nation's Report Card: The creation and evolution of the National Assessment Governing Board (NAGB). Published by National Assessment Governing Board, U. S. Department of Education.

Wagner, Lon (1995). Inside Total Quality Management. The Virginian-Pilot. Published by Landmark Communications, Inc.

Wellman, Jane V. (2001). Assessing state accountability systems. Change. March/April 2001.

Whitaker, B. (2000). Business groups that back superintendent spend big in San Diego school board race. New York Times Oct. 25, 2000. p.A16.

White House, The (2001). Brief on Educational Progress 1992-2000. Retrieved Feb. 2001 from http://www.whitehouse.gov/WH/New/html/edprogress_reports.html

Wilgoren, J. (2000). Young Blacks turn to vouchers as Civil Rights issue. The New York Times. Oct. 9, 2000. p.1, A18.

Yardley, T. (2000). Study casts doubt on Texas test scores, and gives the Democrats ammunition. New York Times (Oct. 25, 2000). p.A21

VITA

Michael Duane Bradshaw was born in Bermuda as the eldest brother to Derek, Kevin, Judy, Laurie and Troy and son to Daphne Anderson and David Saunders. After education at The Central School and then The Berkeley Institute, in 1969 – 1972 he went overseas to England to earn a B.Sc. Hons. in Biochemistry and Genetics from Leeds University. In 1976 he completed a Masters of Philosophy – Developmental Biology - at the University of Sussex. In 1983 he began a service of more than 10 years in the Bermuda Regiment and retired eventually with the rank of Captain.

Michael's varied work experience in education includes experience as a science tutor at a private tutorial institution in England (1974 – 1976); Biology master at an all girls senior science secondary school in West Africa (1976 – 1981); science mentor at Milton Margai Teachers Training College (1977 – 1981) and at Sierra Leone National School of Nursing (1979 – 1981). At Bermuda College (1982 – 2002) he has been senior lecturer (Biology); Chair of the Dept. of Natural Sciences; Associate Dean of the Faculty of Applied Science and Technology and now is Coordinator for Planning and Accreditation.

Educational interests include Membership of the Management Committee of the Berkeley Education Society (A century old private society that founded a school to provide quality education independent of being white or male) and Research Member for Phi Delta Kappa International (Bermuda Chapter) – a professional fraternity for educators. His lifelong educational philosophy of empowerment states that ignorance shackles and limits the human potential. Conversely, regardless of any descriptor of gender or ethnicity or any bias imaginable, it is education that shall make one truly free.

Permanent address: 5A Westdale Lane, Southampton , WK 01 Bermuda College

This dissertation was typed by Michael Duane Bradshaw with assistance from his wife, Janet Smith Bradshaw.